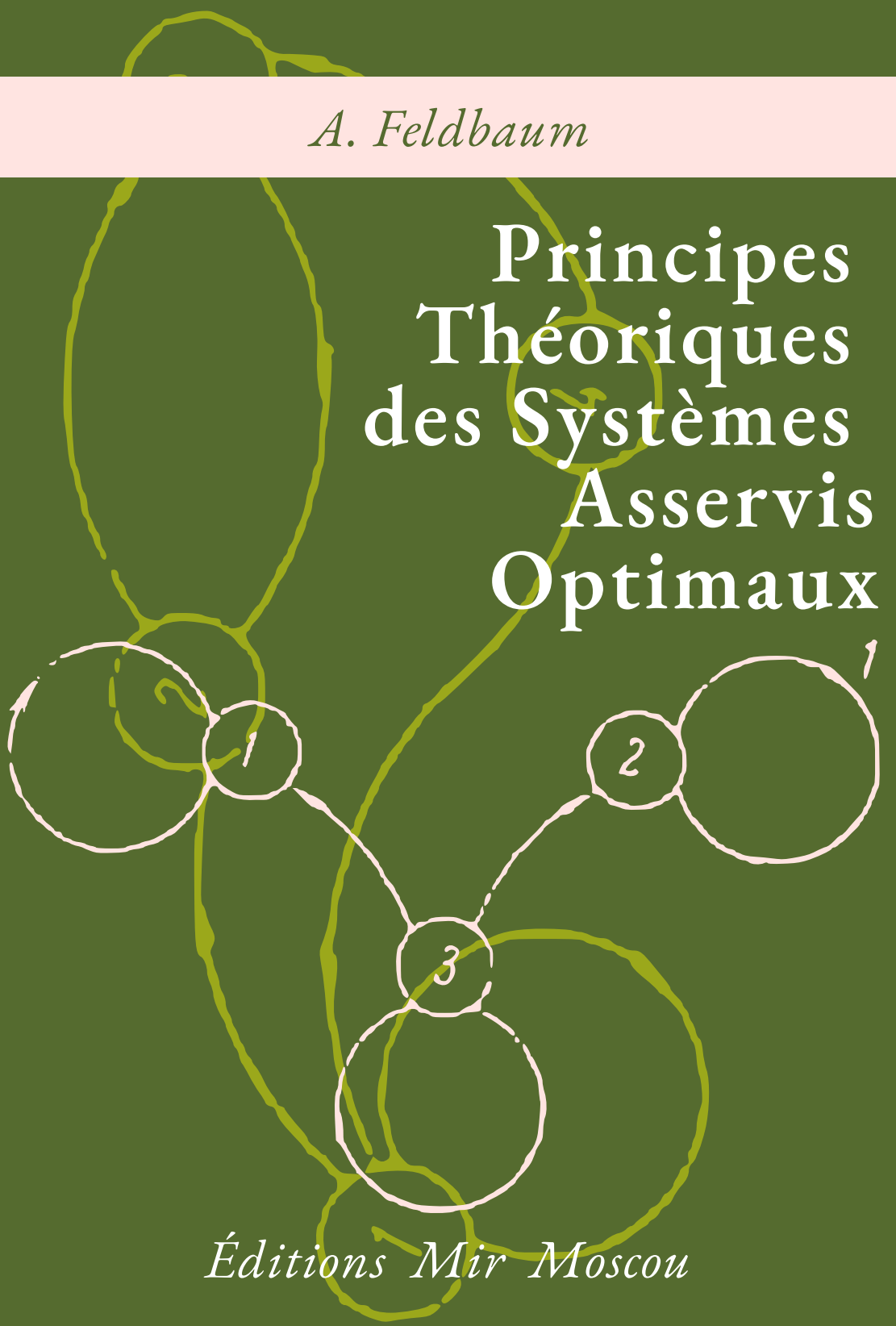
A. Feldbaum

Principes Théoriques des Systèmes Asservis Optimaux



Éditions Mir Moscou

А. А. ФЕЛЬДБАУМ

ОСНОВЫ ТЕОРИИ ОПТИМАЛЬНЫХ АВТОМАТИЧЕСКИХ СИСТЕМ

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

PRINCIPES THÉORIQUES DES SYSTÈMES ASSERVIS OPTIMAUX

PAR

A. FELDBAUM

ÉDITIONS MIR · MOSCOU

CDU 62.50 (075.8) = 40

Traduit du russe
par V. Polonski

На французском языке

© Traduction française Editions Mir
U.R.S.S. 1973

Ф $\frac{3313-262}{041(01)-73}$

TABLE DES MATIÈRES

Avant-propos 7

Chapitre premier. PROBLÈME DU SYSTÈME OPTIMAL 9
§ 1. Intérêt de la théorie des systèmes optimaux . . . 9
§ 2. Classification des systèmes optimaux 13
§ 3. Critères d'optimalité 20
§ 4. Introduction de l'information sur l'objet commandé dans un organe de commande 26
§ 5. Formulation des problèmes de la théorie des systèmes optimaux 32

Chapitre II. MÉTHODES MATHÉMATIQUES APPLIQUÉES À LA THÉORIE DES SYSTÈMES OPTIMAUX 39
§ 1. Eléments de calcul des probabilités 39
§ 2. Méthodes variationnelles 58
§ 3. Programmation dynamique 72
§ 4. Principe du maximum 84

Chapitre III. SYSTÈMES OPTIMAUX À INFORMATION COMPLÈTE SUR L'OBJET COMMANDÉ 109
§ 1. Problème de commande en temps minimal; méthode de l'espace de phase 109
§ 2. Application des méthodes variationnelles classiques . . 136
§ 3. Application de la programmation dynamique 153
§ 4. Application du principe du maximum 171

Chapitre IV. SYSTÈMES OPTIMAUX À INFORMATION INCOMPLÈTE MAXIMALE SUR L'OBJET COMMANDÉ 183
§ 1. Systèmes continus à information maximale sur l'objet 183
§ 2. Systèmes continus à états discrets et systèmes discrets à information maximale sur l'objet 199

Chapitre V. SYSTÈMES OPTIMAUX À ACCUMULATION INDÉPENDANTE (PASSIVE) DE L'INFORMATION SUR L'OBJET 222
§ 1. Principaux problèmes de la théorie des systèmes optimaux à accumulation indépendante de l'information 222
§ 2. Eléments de la théorie générale des décisions statistiques . . 240

§ 3. Application de la théorie des décisions statistiques aux systèmes de commande automatique 251

Chapitre VI. SYSTÈMES OPTIMAUX À ACCUMULATION ACTIVE DE L'INFORMATION 271

§ 1. Formulation d'un problème simple relatif à un système optimal de commande duale 271
§ 2. Résolution du problème et exemples les plus simples 282
§ 3. Exemples des systèmes irréductibles 297
§ 4. Généralisation aux systèmes avec inertie 310
§ 5. Généralisation aux systèmes markoviens 330
§ 6. Schémas fonctionnels des organes de commande optimaux . . . 344

Chapitre VII. CERTAINES APPLICATIONS DE LA THÉORIE DE LA COMMANDE DUALE 364
§ 1. Synthèse d'un modèle auto-adaptatif optimal 364
§ 2. Commande optimale d'un système à retard pur 383
§ 3. Commande avec facteurs indirects 392
§ 4. Systèmes d'auto-apprentissage et commande duale 405

CONCLUSION 421

BIBLIOGRAPHIE 427

INDEX . . . 441

AVANT-PROPOS À L'ÉDITION FRANÇAISE

Le problème d'optimalisation est aujourd'hui le problème clef de la théorie de la commande dans les domaines technique, économique et social.

Il n'y a donc rien d'étonnant que ces dernières années on a vu paraître de nombreux ouvrages traitant des divers aspects de ce problème ainsi que des différentes tentatives de leur résolution.

Il manquait toutefois un livre qui présenterait sous une optique commune l'ensemble de divers approches et problèmes que pose la théorie des systèmes optimaux. Dans une large mesure cette lacune a été comblée par la publication en russe des « Principes théoriques des systèmes asservis optimaux » d'Alexandre Feldbaum. Le présent ouvrage constitue la traduction des « Principes », spécialement remaniée à l'intention du lecteur français.

Dans un ouvrage relativement restreint l'auteur a réussi à brosser un tableau assez complet de l'état actuel de la commande optimale. On y trouve les éléments de mathématiques nécessaires, sans qu'il soit besoin de recourir à des ouvrages spécialisés, et l'exposé de la théorie générale des systèmes optimaux à l'information a priori complète et incomplète, avec accumulation active et passive de l'information.

A. Feldbaum a également développé dans son livre la théorie de la commande duale qui attire ces derniers temps l'attention des spécialistes.

Un exposé précis, clair et imagé, accompagné de nombreux exemples rend ce livre accessible aux étudiants et aux jeunes ingénieurs se spécialisant dans le domaine de la cybernétique, tandis que les résultats originaux et l'approche générale du problème peuvent présenter un intérêt certain pour les chercheurs scientifiques.

Nous espérons que cette traduction française sera bien reçue par les étudiants, ingénieurs et chercheurs francophones.

Iakov Tsypkine

Moscou 1972

CHAPITRE PREMIER

PROBLÈME DU SYSTÈME OPTIMAL

§ 1. Intérêt de la théorie des systèmes optimaux

La technique de transmission et de traitement de l'information, dite technique cybernétique, a connu ces dernières années un développement rapide. L'essor de l'automatique, sa branche maîtresse, est caractérisé par une extension rapide des systèmes asservis, par l'application de ces derniers dans les domaines de plus en plus nombreux. Mais l'essor de l'automatique n'est pas seulement « superficiel », il va aussi « en profondeur ». On voit apparaître de nouveaux principes de commande automatique, de nouveaux types de systèmes asservis aptes à résoudre des problèmes de gestion de plus en plus compliqués et à remplacer l'homme dans des sphères de son activité toujours plus importantes.

Les systèmes automatiques deviennent de plus en plus compliqués. Dans les régulateurs automatiques simples des premiers types, les organes de mesure, de commande et d'exécution étaient souvent réunis. Dans les systèmes de commande automatique modernes ces organes sont dans la plupart des cas des dispositifs particuliers, parfois très complexes. Souvent l'unité centrale du système, l'organe de commande, représente un calculateur analogique ou digital, muni d'une loi de commande dite encore algorithme. Les machines de commande actuelles, universelles ou spécialisées, permettent de mettre en œuvre les algorithmes les plus complexes. Les vitesses de calcul réalisables avec les machines modernes s'élèvent à des centaines de milliers et à des millions d'opérations élémentaires par seconde. C'est pourquoi pour de nombreuses applications l'organe de commande peut être considéré comme un dispositif non inertiel.

Des tendances analogues se manifestent également dans d'autres branches de la cybernétique. Les communications à de très grandes distances, la télémétrie et la télécommande, la séparation des images radar et de télévision des bruits de fond et de parasites artificiels, la reconnaissance des images, autant de problèmes qui imposent la réalisation des algorithmes complexes du traitement de l'information. Un exemple représentatif de tels systèmes est fourni par les dispositifs d'identification des images. Les dispositifs de ce type

ne s'emploient pour le moment que pour identifier des lettres et des chiffres imprimés ou écrits à la main, ainsi que des commandes données à la machine de vive voix. Or, ces dispositifs sont très prometteurs en perspective; c'est l'identification et la classification des images et des situations complexes brouillées par l'ambiance, les bruits, etc., qui permettent à l'homme de s'adapter facilement au milieu ambiant et de prendre des décisions correctes et justifiées. Les algorithmes susceptibles de simuler les fonctions relativement complexes de l'intelligence humaine sont nécessairement complexes; l'organe appelé à les appliquer doit donc être une machine à calculer suffisamment complexe elle aussi.

L'essor de la technique va de pair avec celui de la théorie générale de la commande, qui est à la base d'un grand nombre de branches techniques associées à la transmission et au traitement de l'information. Cette théorie générale s'appelle cybernétique technique; cette dernière fait partie de la cybernétique générale (cf. [1.1]), dont l'objet est l'étude des processus de commande et de traitement de l'information; ces processus sont caractéristiques des systèmes techniques, mais aussi des êtres vivants et des collectivités constituées d'êtres vivants ou de machines.

La cybernétique générale, de même que la cybernétique technique, tire son origine d'un long processus de généralisation des théories, idées et principes disparates établis par des disciplines particulières. Ce processus de généralisation se poursuit encore aujourd'hui. Jusqu'à ces derniers temps la théorie de la commande se développait dans plusieurs directions peu liées entre elles. Ce n'est qu'à présent qu'une tendance se manifeste pour rapprocher ces directions diverses et préciser les notions, idées, méthodes et théories générales.

On distingue dans la cybernétique technique actuelle les principales branches suivantes. L'une des plus importantes est celle qui concerne les principes et la théorie de l'organisation des systèmes de commande complexes, composés d'un grand nombre d'éléments, dont les interactions sont compliquées et les conditions de travail difficiles.

Une autre branche importante de la cybernétique technique s'occupe de la théorie et des principes de fonctionnement des systèmes auto-adaptatifs. Le processus d'adaptation automatique consiste à modifier les paramètres, les caractéristiques et, en général, les propriétés du système ou de ses parties, en agissant soit sur une boucle ouverte, soit par régulation automatique, soit encore par recherche automatique. Ainsi, dans un système d'optimisation automatique l'optimisateur modifie par recherche automatique la caractéristique de l'organe de commande de façon à assurer son meilleur fonctionnement possible lors du changement des propriétés de l'objet gouverné. C'est dans cette direction qu'évolue la théorie des systèmes d'auto-adaptation et d'auto-apprentissage.

L'objet du présent ouvrage est la troisième branche principale de la cybernétique technique, notamment la théorie des processus optimaux (meilleurs dans un certain sens), la théorie des systèmes de commande optimaux, ainsi que des systèmes optimaux de transmission et de traitement de l'information. De nos jours les problèmes associés aux systèmes optimaux deviennent cardinaux et voici pourquoi :

a) Tout système scientifiquement justifié est optimal du fait qu'en le choisissant nous le préférons par là même à d'autres ; nous considérons par conséquent que dans un certain sens il vaut mieux que les autres. Les critères qui président à ce choix (nous dirons dans la suite que ce sont les critères d'optimalité) varient suivant les cas. Mais quel que soit le choix, on fait toujours appel à un critère d'optimalité. S'il n'en était pas ainsi, il serait impossible de faire un choix justifié.

Autrefois, lorsque la production n'était pas encore automatisée et était basée sur l'expérience et les acquis humains, lorsque les techniques de mesure et de calcul étaient encore peu évoluées, les tentatives d'interpréter et de définir les critères d'optimalité et d'autant plus de construire des systèmes optimaux étaient dans la plupart des cas vouées à l'échec. Nous vivons une époque nouvelle, époque marquée par l'automatisation poussée des processus industriels. C'est pourquoi l'intérêt des problèmes de la commande optimale croît sans cesse. Leur résolution permettra d'élever l'efficacité et le rendement des équipements industriels, d'améliorer la qualité de la production, de réduire la consommation d'énergie et des matières premières, etc.

b) Une loi de la nature est une proposition de caractère restrictif délimitant dans un certain domaine le possible de l'impossible. Les lois générales de la cybernétique doivent également permettre de dire ce qu'on peut ou ne peut pas faire dans telles ou telles conditions réelles. Elles peuvent donc être énoncées sous la forme de propositions sur le « plafond des possibilités » de la commande. Or, la tâche de la théorie des systèmes optimaux consiste précisément à rechercher ce « plafond ».

Par conséquent, la théorie des systèmes optimaux doit notablement contribuer à l'établissement des lois générales de la cybernétique. Bien que ce soit dans l'avenir, il semble que cet avenir ne se fera pas beaucoup attendre.

Les problèmes relatifs aux systèmes optimaux intéressent de nombreux domaines de la technique cybernétique. Ce sont notamment les problèmes de commande en temps minimal ; de filtrage optimal du signal mélangé avec un bruit ; de construction des détecteurs de signaux optimaux, des appareils de « prévision » optimaux ; d'établissement des méthodes d'identification des images, de la stratégie optimale de la recherche automatique, etc. Pourtant, tous

ces problèmes, quoique bien différents à première vue, sont liés par de profonds liens internes. Dans ce qui suit nous allons insister sur les traits communs et les rendre évidents, car ce sont précisément ces traits qui sont à la base de la théorie unique des systèmes optimaux.

En plus des branches énumérées ci-dessus on pourrait en citer d'autres, tout aussi importantes, par exemple l'étude de la stabilité et des auto-oscillations des systèmes. En outre, les branches aussi fondamentales que la théorie générale des systèmes (dont fait partie la théorie des automates) et la théorie générale des signaux (dont fait partie la théorie de l'information) appartiennent à la cybernétique générale au moins pour autant qu'à la cybernétique technique.

Ces branches ne sont nullement indépendantes les unes des autres. Ainsi, la théorie des automates finis étudie le problème très intéressant qui consiste à définir la structure la plus simple d'un automate capable de réaliser l'algorithme imposé. Un tel automate sera optimal par sa simplicité. Parmi les problèmes à résoudre il y a également l'étude des automates à fiabilité maximale pour la complexité donnée, c'est-à-dire des systèmes optimaux par la fiabilité. Les questions relatives à l'adaptation automatique s'associent encore plus intimement à la théorie des systèmes optimaux. La majorité de ces questions est traitée par la théorie des systèmes d'optimisation automatique liée à la théorie des systèmes optimaux par de nombreuses attaches dont voici quelques-unes.

a) Un système optimal est l'idéal auquel tend (sans l'atteindre parfois) un système d'optimisation automatique.

b) Lorsque les caractéristiques de l'objet gouverné varient assez lentement, l'organe de commande primaire peut se construire d'après la théorie des systèmes optimaux, mais il faut prévoir la variation de ces paramètres. L'organe de commande secondaire, constitué d'un optimalisateur automatique, en observant le fonctionnement du système, modifie les paramètres du primaire de façon à assurer que le système reste optimal malgré les modifications imprévues des caractéristiques de l'objet gouverné.

c) Lorsque les caractéristiques de l'objet varient assez rapidement, les performances d'un système fonctionnant, comme indiqué en b), peuvent s'écarter sensiblement des performances optimales. Alors le problème se pose de rechercher le meilleur algorithme pour trouver les méthodes de commande optimales, de construire un système optimal de la recherche automatique ou, en général, un système optimal d'adaptation automatique.

La théorie des systèmes optimaux est étroitement liée à d'autres branches de la cybernétique technique ne serait-ce que parce que toute tâche assumée par un organe de commande peut être réalisée de la meilleure façon, optimale dans un certain sens. Par conséquent, dans tout domaine de la cybernétique technique, on rencontre des problèmes de construction des systèmes optimaux.

Nous avons souligné plus haut la valeur purement théorique de la théorie des systèmes optimaux. Pour ce qui est de son intérêt pratique, elle présente deux aspects bien différents. Premièrement, il est impossible de construire sans elle des systèmes asservis optimaux ou quasi optimaux. En effet, même dans les cas les plus simples, l'intuition d'un ingénieur ne peut suffire à dégager les lois optimales qui régissent le fonctionnement des organes de commande. Or, l'application de ces lois aux cas concrets du fonctionnement des machines les plus différentes peut donner un effet économique très important ; les critères applicables dans le domaine des techniques militaires sont d'un tout autre type, mais là aussi l'intérêt des systèmes optimaux est évident. Deuxièmement, la théorie des systèmes optimaux permet d'évaluer le « plafond » qu'on peut atteindre dans un système optimal et de comparer ce « plafond » avec les caractéristiques d'un système déjà réalisé, non optimal. Cette comparaison permet de juger s'il faut chercher un système optimal ou bien se contenter du système existant.

§ 2. Classification des systèmes optimaux

La figure 1.1 représente le schéma fonctionnel d'un système de commande automatique. La lettre A désigne l'organe de commande et la lettre B l'objet gouverné. A la sortie de l'objet B apparaît

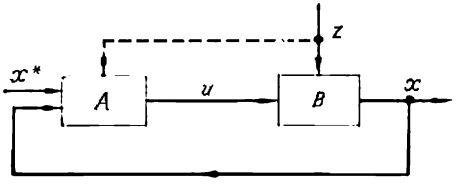


Fig. 1.1.

la *grandeur commandée* x. Par grandeur commandée on entend les paramètres caractéristiques de l'objet gouverné. Dans le cas général il existe plusieurs paramètres de ce type : $x_1, \ldots, x_n$. Il est commode de les considérer comme les coordonnées du vecteur $\bar{x}$:

$$\bar{x} = (x_1, \ldots, x_n). \tag{1.1}$$

Le vecteur $\bar{x}$ s'appelle également vecteur de sortie ou grandeur de sortie de l'objet B.

L'organe de commande A envoie sur l'entrée de l'objet B la commande u. Plusieurs commandes $u_1, u_2, \ldots, u_r$ peuvent être considérées comme les coordonnées u_j ($j = 1, \ldots, r$) du vecteur $\bar{u}$:

$$\bar{u} = (u_1, \ldots, u_r). \tag{1.2}$$

A l'entrée de l'organe de commande A on amène la *consigne* x^* qui est une instruction définissant la grandeur de sortie x de l'objet B. Cette instruction doit concrétiser le but de la commande (le concept correspondant est précisé au paragraphe suivant). Elle peut être constituée d'une collection de n grandeurs $x_1^*, \ldots, x_n^*$ que nous allons considérer comme les coordonnées du vecteur $\bar{x}^*$:

$$\bar{x}^* = (x_1^*, \ldots, x_n^*). \tag{1.3}$$

On peut exiger, par exemple, que le cas idéal vérifie les conditions

$$x_i = x_i^* \quad (i = 1, \ldots, n), \tag{1.4}$$

où x_i^* sont les fonctions du temps données.

Les systèmes de commande automatique forment deux classes : systèmes en boucle ouverte et systèmes en boucle fermée, ces derniers étant également dits systèmes à réaction. Dans les systèmes en boucle ouverte, l'organe de commande A ne reçoit pas d'information sur l'état réel $\bar{x}$ de l'objet B. Dans les systèmes en boucle fermée, l'organe A reçoit cette information par la boucle de réaction (en bas de la figure 1.1). Le principe du fonctionnement d'un système de commande en boucle fermée peut être caractérisé brièvement de la façon suivante : si la grandeur $\bar{x}$ ne correspond pas aux exigences $\bar{x}^*$, l'organe de commande A exerce sur l'objet B une action $\bar{u}$ qui rapproche $\bar{x}$ de $\bar{x}^*$.

L'écart entre la grandeur $\bar{x}$ et les consignes peut être dû à des causes diverses.

a) Utilisation incorrecte, imprécise ou retardée par l'organe A de l'information qu'il contient ou qu'il reçoit sur les caractéristiques et l'état de l'objet et le but de la commande.

b) Possibilités limitées de la commande, c'est-à-dire l'impossibilité pour telle ou telle raison d'amener à l'objet B des signaux de commande $\bar{u}$ pouvant assurer à l'objet le comportement imposé $\bar{x}$. En pratique, les possibilités de commande sont toujours limitées et ceci ne doit jamais être perdu de vue.

c) Il se peut que l'écart entre $\bar{x}$ et la valeur de consigne soit dû à une *perturbation* z imprévisible et incontrôlable, agissant sur l'objet B et intervenant dans sa grandeur de sortie $\bar{x}$. Si les différentes parties de l'objet B subissent des perturbations $z_1, \ldots, z_l$, on peut les représenter sous la forme du vecteur $\bar{z}$:

$$\bar{z} = (z_1, \ldots, z_l). \tag{1.5}$$

Les perturbations que subit l'objet gouverné B peuvent provoquer une modification imprévue de ses caractéristiques. L'action exercée

sur l'objet par la modification de la charge peut être considérée comme un cas particulier de la manifestation d'une perturbation.

Supposons que l'algorithme de l'organe de commande A assure un bon fonctionnement du système pour des caractéristiques définies de l'objet B. La variation de ces caractéristiques peut altérer le fonctionnement du système et entraîner des écarts importants entre la grandeur $\bar{x}$ et la valeur de consigne.

Dans de nombreux cas, le principe de la réaction rend possible l'observation des conditions imposées à la grandeur $\bar{x}$ même en présence d'une importante perturbation $\bar{z}$ subie par l'objet B. Pourtant, lorsque les caractéristiques de l'objet B sont complexes et varient rapidement dans une large plage, le problème de commande devient plus difficile. L'information sur la perturbation $\bar{z}$, ne serait-ce que sur certaines de ses composantes $z_1, \ldots, z_{l'}$ $(l' < l)$, peut contribuer grandement à la résolution du problème et améliorer le résultat de la commande. Supposons que la perturbation soit mesurée et que les résultats des mesures soient canalisés (cf. trait interrompu de la figure 1.1) jusqu'à l'organe de commande A. Ce dernier peut alors calculer et fournir une commande $\bar{u}$ telle qu'elle neutralise la perturbation $\bar{z}$ et fait prendre à la grandeur de sortie $\bar{x}$ de l'objet B une valeur qui correspond mieux à celle de consigne. Cette méthode est dite de compensation.

Dans plusieurs cas, lorsqu'il est assez simple de mesurer l'action perturbatrice principale, la méthode de compensation ou sa combinaison avec le principe de réaction présente de grands avantages.

Dans la plupart des cas les propriétés de l'objet B sont imposées et ne doivent pas être modifiées. Par contre, l'algorithme de l'organe de commande A n'étant pas défini d'avance, on peut le choisir dans une classe très étendue des solutions possibles *). Le problème de construction d'un système optimal se ramène ainsi à la recherche d'un organe de commande A tel que dans un certain sens il soit capable de commander l'objet B de la meilleure façon.

Si A ne doit répondre à aucune condition supplémentaire, l'algorithme de l'organe optimal A est défini par les facteurs suivants, associés à l'objet B et au mode de sa liaison avec A :

1) caractéristiques de l'objet B ;
2) conditions qui lui sont imposées (but de la commande) ;
3) information sur l'objet B fournie à l'organe de commande A.

Une position correcte du problème exige une étude détaillée de ces facteurs. On peut représenter chacun de ces facteurs par une

---

*) Il arrive souvent que l'organe de commande comporte une partie de puissance invariable; il faut alors rapporter cette dernière à l'objet gouverné et la considérer comme l'une de ses parties. C'est pourquoi le concept d'« objet gouverné » est remplacé parfois par celui de « partie invariable du système ».

certaine direction, orthogonale aux autres (fig. 1.2), et mettre en correspondance à chaque type de systèmes optimaux un point ou un domaine de l'espace tridimensionnel ainsi obtenu.

Les directions de la figure 1.2 traduisent les branches de la classification des systèmes optimaux. L'avantage que présente une telle classification est qu'elle permet de définir correctement la place de chaque type de systèmes optimaux parmi les autres types. L'étude de tous les types possibles de systèmes optimaux considérés sous une optique générale révèle l'unicité des lois fondamentales de la théorie, indépendamment des différences qui existent entre les types particuliers des systèmes.

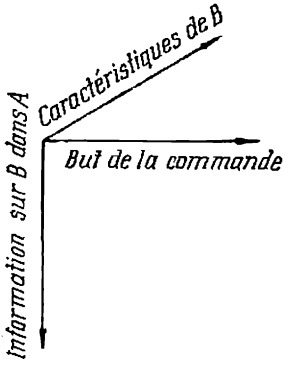


Fig. 1.2.

La première branche de la figure 1.2 correspond à la classification d'après les caractéristiques des objets. La figure 1.1 montre que l'objet B est caractérisé par la relation entre sa grandeur de sortie $\bar{x}$ et les grandeurs d'entrée $\bar{u}$ et $\bar{z}$, qui peut s'écrire

$$\bar{x} = \bar{F}(\bar{u}, \bar{z}). \tag{1.6}$$

Dans le cas général, la relation $\bar{F}$ est un opérateur, c'est-à-dire une loi de correspondance entre deux ensembles de fonctions. Par exemple, dans la formule (1.6), le vecteur fonction $\bar{x}$ est défini par le type des vecteurs fonctions $\bar{u}$ et $\bar{z}$. L'opérateur $\bar{F}$ de l'objet peut être défini de façons différentes: par des formules, des graphiques ou des tableaux. Souvent cette relation est donnée sous la forme d'un système d'équations différentielles, comme par exemple

$$\left.\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ z_1, \ldots, z_l;\ t), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dx_n}{dt} &= f_n(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ z_1, \ldots, z_l;\ t). \end{aligned}\right\} \tag{1.7}$$

Dans le cas général f_i sont des fonctions non linéaires de $x_1, \ldots, x_n$; $u_1, \ldots, u_r$; $z_1, \ldots, z_l$ et du temps t. Avec les notations vectorielles

$$\frac{d\bar{x}}{dt} = \left(\frac{dx_1}{dt}, \ldots, \frac{dx_n}{dt}\right); \quad \bar{f} = (f_1, \ldots, f_n) \tag{1.8}$$

on peut écrire les équations (1.7) sous une forme vectorielle condensée

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}, \bar{u}, \bar{z}, t). \tag{1.9}$$

Dans cette expression le vecteur $\bar{f}$ est une fonction des vecteurs $\bar{x}$, $\bar{u}$, $\bar{z}$ et du scalaire t. Si le vecteur fonction $\bar{u}(t)$ est connu et la perturbation $\bar{z}(t)$ imposée, les équations (1.7) ou (1.9) permettent de trouver le vecteur $\bar{x}(t)$ pour les conditions initiales données, c'est-à-dire lorsque l'on connaît le vecteur $\bar{x}^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$, où

$$x_i^{(0)} = (x_i)_{t=0} \qquad (i = 1, \ldots, n). \tag{1.10}$$

Les modes de classification des opérateurs des objets peuvent être les plus différents. Nous ne passerons en revue que les plus importants. Considérons d'abord la partition des systèmes en systèmes continus, continus à états discrets et discrets. Dans les systèmes du premier type les grandeurs sont considérées à n'importe quel instant, de plus, ces grandeurs peuvent varier en permanence et, en principe, leur niveau peut être quelconque (fig. 1.3). Ainsi, d'après les définitions adoptées dans les techniques de communication et de régulation automatique, ces grandeurs ne sont pas quantifiées ni dans le temps ni en amplitude. Telles sont, par exemple, les solutions $x_1, \ldots, x_n$ des équations (1.7) qui sont fonctions du temps continu t. Dans les systèmes du second type, utilisant par exemple la commande par impulsions ou numérique, ainsi que la modulation d'impulsions pour la transmission des signaux, les valeurs des grandeurs ne présentent un intérêt qu'à des instants discrets $t = t_0, t_1, t_2, \ldots$ Si, dans ces conditions, les niveaux acceptables des grandeurs sont quelconques, cela signifie que ces dernières sont quantifiées dans le temps et non pas en amplitude. Les systèmes respectifs s'appellent systèmes continus à états discrets. L'opérateur d'un système continu à état discret peut être donné, par exemple, par des équations aux différences finies. Désignons par $x_i(m)$ la valeur de x_i à l'instant $t = t_m$ (cf. fig. 1.3). Désignons ensuite par

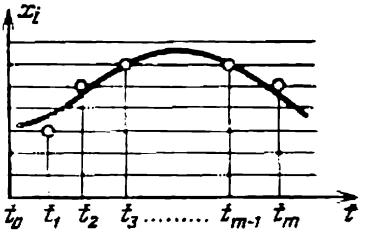


Fig. 1.3.

$$\bar{x}(m) = [x_1(m), \ldots, x_n(m)] \tag{1.11}$$

le vecteur $\bar{x}$ à l'instant $t = t_m$ et par $\bar{u}(m)$ le vecteur $\bar{u}$ à ce même instant $t = t_m$. Les équations aux différences finies qui associent les valeurs futures $x_i(m+1)$ aux valeurs antérieures $x_i(m)$ peuvent s'écrire

$$x_i(m+1) = g_i[x_1(m), \ldots, x_n(m); u_1(m), \ldots, u_r(m); z_1(m), \ldots, z_l(m); m] \quad (i = 1, \ldots, n), \tag{1.12}$$

où g_i sont en général des fonctions non linéaires de leurs arguments. Les équations (1.12) peuvent être mises sous une forme vectorielle

$$\bar{x}(m+1) = \bar{g}[\bar{x}(m)\,;\ \bar{u}(m)\,;\ \bar{z}(m)\,;\ m], \tag{1.13}$$

$\bar{g}$ étant ici le vecteur aux composantes $g_1, \ldots, g_n$.

Dans le troisième type des systèmes ne sont autorisés que des niveaux discrets définis des grandeurs considérées. Les systèmes dans lesquels les grandeurs sont quantifiées aussi bien dans le temps qu'en amplitude sont dits discrets (réseau de niveaux sur la figure 1.3). L'opérateur d'un système discret peut être caractérisé, par exemple, par les équations (1.12); toutefois, toutes les grandeurs figurant dans ces équations doivent avoir des niveaux autorisés. Les fonctions g_i, notamment, ne peuvent prendre que les valeurs correspondant aux valeurs autorisées pour x_i.

Un grand nombre d'ouvrages est consacré aux systèmes continus optimaux (cf. par exemple [3.2, 3.14, 3.16-3.18, 3.21-3.23]).

Les systèmes continus à états discrets optimaux ont fait l'objet de [3.19, 3.20, 3.24, 3.30, 3.31], et les systèmes discrets optimaux, de [4.12].

Les objets peuvent être classés également d'après le type de leurs équations. Dans la plupart des travaux consacrés aux systèmes optimaux sont étudiés les objets aux paramètres localisés, dont la variation est décrite par des équations différentielles ordinaires. Pourtant, dans [3.42] on trouve la position du problème et dans [3.43, 3.44] sa résolution dans le cas des objets aux paramètres répartis, décrits par des équations aux dérivées partielles et des équations intégrales.

Parmi les caractéristiques de l'objet B on compte également les *contraintes* de types variés. Par exemple, les composantes $u_1, \ldots, u_r$ du vecteur $\bar{u}$ (cf. formule 1.2) ne peuvent pas avoir des valeurs quelconques. Elles ne peuvent pas dépasser certaines limites par suite des propriétés physiques de l'objet ou ne le doivent pas pour ne pas compromettre le fonctionnement normal de l'objet.

Souvent les contraintes sont écrites sous la forme

$$|u_1| \leqslant U_1, \ \ldots, \ |u_r| \leqslant U_r, \tag{1.14}$$

où $U_1, \ldots, U_r$ sont des constantes données. Il se peut que la limitation porte sur une fonction de plusieurs actions de commande, comme, par exemple, dans le cas

$$\sum_{\nu=1}^{r} \lambda_\nu^2 u_\nu^2 \leqslant N, \tag{1.15}$$

où λ_ν^2 et N sont des constantes ou des fonctions du temps données.

Considérons l'espace de dimension r du vecteur $\bar{u}$ aux coordonnées cartésiennes $u_1, \ldots, u_r$. Les conditions (1.14) ou (1.15) sont des

cas particuliers limitant la disposition de l'extrémité du vecteur $\bar{u}$ à un certain domaine admissible $\Omega(\bar{u})$ de cet espace. L'expression « $\bar{u}$ appartient au domaine $\Omega(\bar{u})$ » s'écrit sous une forme symbolique

$$\bar{u} \in \Omega(\bar{u}). \tag{1.16}$$

Dans le cas particulier des conditions (1.14) le vecteur $\bar{u}$ est borné par un parallélépipède de dimension r, et dans le cas des conditions (1.15) par un ellipsoïde de dimension r dans l'espace $\bar{u}$.

Les contraintes peuvent être imposées non seulement aux commandes u_j, mais encore aux coordonnées x_i $(i = 1, \ldots, n)$ de l'objet B. Ainsi, certaines fonctions ou fonctionnelles $H_\mu(\bar{x})$ de ces coordonnées ne doivent pas dépasser certaines limites qu'on peut prendre égales à zéro sans perte de généralité :

$$H_\mu(x_1, \ldots, x_n) = H_\mu(\bar{x}) \leqslant 0 \quad (\mu = 1, \ldots, m). \tag{1.17}$$

Les fonctions ou fonctionnelles $H_\mu(\bar{x})$ peuvent être considérées comme les coordonnées d'un vecteur $\bar{H}(\bar{x})$ à m dimensions. Les conditions (1.17) imposent des contraintes à la disposition de ce vecteur. Si H_μ sont des fonctions univalentes de $\bar{x}$, les conditions (1.17) traduisent le fait que dans l'espace $\bar{x}$ de dimension n le vecteur $\bar{x}$ est également borné par un certain domaine admissible $\Omega(\bar{x})$:

$$\bar{x} \in \Omega(\bar{x}). \tag{1.18}$$

Dans le cas général, les contraintes sont imposées à certaines fonctionnelles L de $\bar{u}(t)$, $\bar{x}(t)$ et $\bar{z}(t)$, c'est-à-dire aux grandeurs dépendant du type des fonctions $\bar{u}$, $\bar{x}$ et $\bar{z}$ dans un intervalle quelconque :

$$L_\mu[\bar{u}(t), \bar{x}(t), \bar{z}(t)] \in \Omega_\mu(L) \quad (\mu = 1, \ldots, m), \tag{1.19}$$

où $\Omega_\mu(L)$ est le domaine admissible de variation de la fonctionnelle L_μ. L'exemple en est fourni par la contrainte du type

$$L = \int_0^T \left[\sum_{\nu=1}^{n} \alpha_\nu x_\nu^2 + \beta u^2\right] dt \leqslant N, \tag{1.20}$$

où T, α_ν, β et N sont des constantes positives *).

Dans les systèmes continus à états discrets et les systèmes discrets les restrictions concernent les grandeurs analogues. On peut rapporter également aux caractéristiques de l'objet B celles de la perturbation z due au milieu ambiant (cf. fig. 1.1 et équations (1.7),

*) La formule de la contrainte peut comporter également le temps t sous une forme explicite.

(1.9)). Parfois on fait entrer les caractéristiques de la perturbation dans l'opérateur de l'objet. Si z_ν sont des fonctions du temps connues ($\nu = 1, \ldots, l$), leurs expressions peuvent être portées dans les équations de l'objet B, où le temps sera présent sous une forme explicite. A des fins méthodologiques il est commode de considérer toute perturbation imprévue z_ν comme appliquée à l'objet de l'extérieur, alors que toute perturbation supposée connue est inclue dans l'opérateur $\overline{F}$.

En s'ajoutant à d'autres actions, à u_j, par exemple, les perturbations z_ν peuvent être amenées aux entrées de l'objet B. Ces actions sont dites additives. Mais les z_ν peuvent agir d'une autre manière, en modifiant les coefficients des équations des circuits ou leurs paramètres. De telles actions sont dites paramétriques. Dans des systèmes non linéaires il est difficile de délimiter ces deux types d'action.

Les perturbations aléatoires z_ν peuvent être des grandeurs ou des processus aléatoires que nous étudierons au chapitre II. Dans le premier cas, au cours d'un processus isolé dans le système, les z_ν peuvent être considérées comme constantes; dans le deuxième cas, les z_ν sont des fonctions aléatoires du temps dont la variation au cours d'un processus ne peut être négligée.

Parfois la présence des perturbations aléatoires z_ν dans les conditions du problème n'est pas explicite. Mais si $\overline{z}$ est aléatoire et $\overline{u}$ donné, d'après (1.7) la grandeur de sortie $\overline{x}$ de l'objet B est un processus aléatoire. Au lieu des caractéristiques de $\overline{z}$ on peut introduire directement les caractéristiques probabilistes conditionnelles du processus $\overline{x}$, fonctions de $\overline{u}$ et des conditions initiales $\overline{x}^0$, donc, donner de cette façon l'opérateur $\overline{F}$ et les caractéristiques de la perturbation aléatoire $\overline{z}$. C'est ainsi qu'est défini l'objet dans [4.12].

§ 3. Critères d'optimalité

La deuxième branche de la classification des systèmes optimaux est basée sur les conditions imposées au comportement de l'objet B (cf. fig. 1.2). L'une des exigences est le *but de la commande*. Quel que soit le cas, on peut considérer que le but de la commande est d'obtenir l'extrémum d'une certaine grandeur Q, dite critère d'optimalité. Suivant les exigences, il s'agit de maximiser ou de minimiser la grandeur Q. Dans le cas général le critère d'optimalité dépend de la consigne $\overline{x}^*$, de même que de la grandeur de sortie $\overline{x}$; il peut également dépendre de $\overline{u}$ et de $\overline{z}$ ainsi que du temps t. Pour fixer les idées, supposons que la grandeur Q est à minimiser :

$$Q(\overline{x}, \overline{x}^*, \overline{u}, \overline{z}, t) = \min. \tag{1.21}$$

Cette condition est une formulation analytique du but de la commande. Notons que Q est une fonctionnelle, c'est-à-dire un nombre défini par le type des fonctions $\bar{x}$, $\bar{x}^*$, $\bar{u}$ et $\bar{z}$. Dans un cas particulier Q s'écrit

$$Q = \int_0^T [x(t) - x^*(t)]^2 \, dt, \tag{1.22}$$

où T est une grandeur fixée. La formule (1.22) montre que la grandeur Q est définie par le type des fonctions $x(t)$ et $x^*(t)$ dans l'intervalle $0 < t < T$.

Dans le choix du critère d'optimalité Q on part des conditions techniques et économiques concrètes, ce qui sort du cadre de la théorie des systèmes optimaux.

La formule (1.21) pour Q permet non seulement de connaître sa valeur minimale possible $Q_{\min}$, mais aussi d'évaluer l'altération du fonctionnement du système en cas d'écart de l'optimum. La mesure de cette altération est donnée par la différence $Q - Q_{\min}$ ou par une fonction monotone quelconque de cette différence qui devient nulle quand $Q = Q_{\min}$.

La classification peut partir des types des critères Q. Ainsi, les critères d'optimalité diffèrent suivant qu'ils se rapportent à un processus transitoire ou stationnaire. Considérons à titre d'exemple les critères intégraux des processus dans des systèmes linéaires. Soit le processus dans un certain système linéaire, décrit par une équation différentielle linéaire à coefficients constants, associant la grandeur d'entrée x^* à celle de sortie x:

$$a_0 \frac{d^n x}{dt^n} + a_1 \frac{d^{n-1} x}{dt^{n-1}} + \ldots + a_n x = b_0 \frac{d^m x^*}{dt^m} + \ldots + b_m x^*. \tag{1.23}$$

On sait que la solution de cette équation s'écrit

$$x(t) = x_s(t) + x_d(t), \tag{1.24}$$

où $x_s(t)$ est la solution particulière de l'équation à membre droit, et $x_d(t)$ est la solution générale de l'équation sans membre droit

$$a_0 \frac{d^n x_d}{dt^n} + a_1 \frac{d^{n-1} x_d}{dt^{n-1}} + \ldots + a_n x_d = 0. \tag{1.25}$$

Donnons l'interprétation physique de la formule (1.24); sous certaines conditions supplémentaires $x_s(t)$ est un processus stationnaire, alors que $x_d(t)$ est un processus transitoire dans le système. Pour un système stable et dans ce qui suit nous supposons qu'il en est ainsi, on a

$$x_d(t) \to 0 \text{ avec } t \to \infty. \tag{1.26}$$

Pour trouver $x_d(t)$ il faut d'abord résoudre l'équation caractéristique du système

$$a_0p^n + a_1p^{n-1} + \ldots + a_n = 0, \tag{1.27}$$

dont les racines sont $p_1, p_2, \ldots, p_n$. Sans perte de généralité, on peut considérer que toutes les racines sont différentes, il vient

$$x_d(t) = C_1e^{p_1t} + \ldots + C_ne^{p_nt}, \tag{1.28}$$

les constantes C_i ($i = 1, \ldots, n$) étant trouvées à partir des conditions initiales

$$\left(\frac{d^kx_d}{dt^k}\right)_{t=0} = \left(\frac{d^kx}{dt^k}\right)_{t=0} - \left(\frac{d^kx_s}{dt^k}\right)_{t=0}. \tag{1.29}$$

Pour se représenter le processus transitoire, il faut construire d'après l'équation (1.28) la courbe de $x_d(t)$. Cependant la solution peut être obtenue d'une manière plus simple en calculant, par exemple, l'intégrale

$$I_1 = \int_0^\infty x_d(t)\,dt. \tag{1.30}$$

Cette intégrale est définie sous sa forme générale comme une fonction des coefficients de l'équation (1.25) et des conditions initiales sans qu'on ait besoin de connaître la fonction $x_d(t)$. Si le signe de $x_d(t)$ ne change pas, par exemple $x_d(t) > 0$ pour un $t \geqslant 0$ quelconque, la décroissance de l'intégrale I_1 correspond en général à l'accélération du processus transitoire. C'est pourquoi I_1 est considéré parfois comme le critère de « qualité » du processus transitoire. Mais pour des processus où $x_d(t)$ change de signe, il se peut qu'une petite valeur de I_1 correspond justement au cas d'un processus à faible amortissement et à allure oscillatoire bien marquée *). C'est pourquoi le domaine d'application du critère I_1 est borné. Dans l'ouvrage [1.3] est proposé un autre critère:

$$I_2 = \int_0^\infty x_d^2(t)\,dt. \tag{1.31}$$

En choisissant convenablement les paramètres ou l'algorithme de l'organe de commande A dans le but de minimiser l'intégrale I_2 on arrive souvent à obtenir une allure satisfaisante du processus transitoire. Le critère I_2 a été appliqué aux systèmes de régulation

*) Il y a des cas, toutefois, où l'on peut appliquer le critère I_1 même lorsque l'allure du processus est nettement oscillatoire. Cf. l'article de A. Voronov dans la revue « Avtomatika i télémékhanika » (« Automatique et télémécanique »), n° 6, 1963.

automatique [1.4, 1.5]. Pourtant il n'est pas rare qu'avec ce critère on obtient des processus transitoires trop oscillatoires. On a donc proposé [1.6] le critère intégral, dit généralisé,

$$I_V = \int_0^\infty V\,dt, \tag{1.32}$$

où V est la forme quadratique des composantes transitoires x_{di} des coordonnées $x_1, \ldots, x_n$ du système :

$$V = \sum_{i,\,j=1}^{n} a_{ij} x_{di} x_{dj}. \tag{1.33}$$

L'interprétation physique du critère intégral généralisé peut être donnée par un exemple très simple dans lequel $x_{d1} = x_1$ est la composante transitoire de l'erreur du système et $\frac{dx_{d1}}{dt} = x_2$. Soit

$$I_V = \int_0^\infty [x_1^2 + T^2 x_2^2]\,dt = \int_0^\infty \left[x_{d1}^2(t) + T^2\left(\frac{dx_{d1}}{dt}\right)^2\right] dt, \tag{1.34}$$

où T = const.

En choisissant les paramètres du système de façon à minimiser l'intégrale I_V, nous interdisons une existence durable des écarts importants x_{d1} (sinon la composante $\int_0^\infty x_{d1}^2\,dt$ de l'intégrale I_V sera trop grande), mais aussi celle des valeurs importantes des dérivées $\frac{dx_{d1}}{dt}$ (sinon c'est la composante $\int_0^\infty \left(\frac{dx_{d1}}{dt}\right)^2 dt$ de l'intégrale I_V qui sera trop grande.) On obtient ainsi un processus transitoire à la fois rapide et progressif, sans oscillations brutales.

Les critères (1.30)-(1.32) s'emploient pour évaluer le processus transitoire $x_d(t)$. Pour apprécier un processus stable $x_s(t)$ on recourt à des critères d'autre type, dont nous donnons un exemple

$$x_{m.q}^2 = \lim_{T\to\infty} \frac{1}{T}\int_0^T x^2(t)\,dt = \lim_{T\to\infty} \frac{1}{T}\int_0^T x_s^2(t)\,dt +$$

$$+ \lim_{T\to\infty} \frac{1}{T}\int_0^T x_d^2(t)\,dt + \lim_{T\to\infty} \frac{2}{T}\int_0^T x_s x_d\,dt. \tag{1.35}$$

D'autres types de critères d'optimalité établis pour des processus transitoires et stationnaires sont décrits dans [1.7-1.9].

Souvent on prend comme critère d'optimalité d'un processus transitoire le temps de régulation ou bien la valeur de l'écart maximal du processus par rapport à une certaine grandeur imposée ou

à une fonction du temps. Dans ce dernier cas on exige pour un système optimal le minimum de l'écart maximal, appelé minimax.

Il importe ici d'insister sur le fait qu'on ne peut pas exiger que *deux* ou plusieurs fonctions d'une ou de plusieurs variables aient simultanément leurs extréma. En effet, les extréma des fonctions ou des fonctionnelles différentes ne correspondent pas en général à la même valeur de l'ensemble des arguments. Donc, il n'existe pas dans le cas général de valeurs des arguments donnant simultanément l'extrémum à deux ou à plusieurs fonctions et fonctionnelles. On peut seulement demander qu'*une* fonction ou fonctionnelle ait son extrémum, tout en imposant à d'autres fonctions et fonctionnelles des conditions de restriction supplémentaires. Ces contraintes peuvent présenter un caractère complexe.

Parfois on recourt à des critères composites. Supposons qu'il faut choisir un vecteur $\bar{x}$ tel qu'il minimise la fonction $Q_1(\bar{x})$, alors que $Q_j(\bar{x}) \leqslant 0$ $(j = 2, \ldots, m)$. Ces dernières inégalités délimitent dans l'espace du vecteur $\bar{x}$ un certain domaine admissible qu'il ne faut pas dépasser. Formellement, les contraintes peuvent être enlevées si l'on fait appel au critère

$$Q(\bar{x}) = Q_1(\bar{x}) + \sum_{j=2}^{m} \beta_j(Q_j)\, Q_j(\bar{x}), \tag{1.36}$$

les fonctions β_j étant ici de la forme

$$\beta_j = \left\{ \begin{array}{r} 0 \text{ avec } Q_j \leqslant 0, \\ \gamma^2 \gg 1 \text{ avec } Q_j > 0 \end{array} \right\} \quad (j = 2, \ldots, m). \tag{1.37}$$

Si le nombre γ^2 est suffisamment grand, le point de minimum de la fonction Q soit se confond avec le minimum Q_1 si ce dernier se trouve à l'intérieur du domaine admissible, soit repose pratiquement sur sa frontière sans la dépasser. On peut construire les fonctions $\beta_j(Q_j)$ également sous la forme de $(1 + Q_j)^{\alpha_j}$, où $\alpha_j \gg 1$. Dans la formule (1.36) le critère est généralement difficile à analyser. Pour éviter les grandes valeurs des coefficients, la formule (1.36) peut être remplacée par la formule suivante:

$$Q(\bar{x}) = \beta_1(Q_2, \ldots, Q_m) \cdot Q_1(\bar{x}) + \sum_{j=2}^{m} \beta_j(Q_j)\, Q_j(\bar{x}), \tag{1.38}$$

où

$$\left. \begin{array}{l} \beta_1(Q_2, \ldots, Q_m) = \left\{ \begin{array}{l} 1, \ Q_j \leqslant 0 \ (j = 2, \ldots, m), \\ 0 \text{ si au moins un des } Q_j > 0; \end{array} \right. \\ \beta_j(Q_j) = \left\{ \begin{array}{l} 1, \ Q_j > 0, \\ 0, \ Q_j \leqslant 0. \end{array} \right. \end{array} \right\} \tag{1.39}$$

Bien que la formule (1.38) devienne alors encore plus compliquée, elle se prête très bien au calcul machine et on l'applique dans certains optimaliseurs automatiques [3.25, 6.6].

Suivant le caractère du critère d'optimalité, on distingue les types suivants des systèmes optimaux :

a) uniformément optimaux ;

b) statistiquement optimaux ;

c) régis par le critère du minimax.

Dans les systèmes du premier type tout processus pris à part est optimal. Ainsi, quelles que soient les conditions initiales ou les consignes (ces dernières devant appartenir à une classe définie d'actions admissibles), dans des systèmes optimaux du point de vue de la rapidité d'action (cf. par exemple, [3.1-3.24]) l'objet est ramené à l'état imposé en un temps minimal.

Dans les systèmes du deuxième type on n'exige pas ou on ne peut pas obtenir un comportement optimal dans chaque processus particulier. Le critère d'optimalité Q a ici une allure statistique. Ces systèmes doivent être optimaux en moyenne. Les critères statistiques s'appliquent aux systèmes qui comportent sous telle ou telle forme des facteurs aléatoires. Donnons un exemple : il s'agit du choix des paramètres $a_1, \ldots, a_k$ d'un organe de commande A dont le schéma est imposé. Supposons que le critère de qualité primaire soit une fonction scalaire quelconque

$$Q_1 = Q_1(a_1, \ldots, a_k; \quad x_1^{(0)}, \ldots, x_n^{(0)}) = Q(\bar{a}, \bar{x}^0), \tag{1.40}$$

$\bar{x}^{(0)}$ étant ici le vecteur des conditions initiales $x_i^{(0)}$ $(i = 1, \ldots, n)$ de l'objet B et $\bar{a}$ le vecteur des paramètres aux coordonnées a_j $(j = 1, \ldots, k)$.

Le critère Q_1 ne peut pas guider directement le choix des paramètres a_j, car des valeurs de $\bar{a}$, optimales pour un type de conditions initiales $\bar{x}^{(0)}$, ne le sont pas en général pour un autre type. Pourtant, si l'on connaît la densité de probabilité a priori $P(\bar{x}^{(0)})$ du vecteur des conditions initiales, on peut prendre comme critère la grandeur Q, l'espérance mathématique, ou comme on l'appelle en physique et en technique, la *valeur moyenne* de la grandeur Q_1. Désignons l'espérance mathématique par M. Alors, comme nous le verrons au chapitre II,

$$Q = M\{Q_1\} = \int\limits_{\Omega(\bar{x}^0)} Q_1(\bar{a}, \bar{x}^0)\, P(\bar{x}^0)\, d\Omega(\bar{x}^0). \tag{1.41}$$

Dans cette formule $\Omega(\bar{x}^{(0)})$ est le domaine de variation du vecteur $\bar{x}^0$, et $d\Omega(\bar{x}^{(0)})$ son élément infinitésimal.

L'interprétation physique de cette estimation est dans le fait que, pour un très grand nombre d'expériences, la grandeur Q se

confond pratiquement avec la moyenne arithmétique des valeurs Q_1 fournies par chacune des expériences.

Dans cette formulation du problème, le système optimal est celui dont les paramètres a_i minimisent la grandeur Q (en général, compte tenu des contraintes supplémentaires).

Dans notre exemple on peut représenter les valeurs initiales aléatoires du processus transitoire comme le résultat de l'action, à l'instant initial, de brèves impulsions aléatoires. Ainsi, on est en présence d'un cas particulier de l'action sur l'objet d'une perturbation aléatoire $\bar{z}$. On peut donner également un autre exemple très simple du système à critère d'optimalité primaire du type $Q_1(\bar{x}, \bar{x}^*, \bar{u}, \bar{z})$, où la perturbation $\bar{z}$ est une variable aléatoire de densité de probabilité $P(\bar{z})$. On peut prendre alors comme critère d'optimalité la valeur moyenne Q de la grandeur Q_1:

$$Q = M\{Q_1\} = \int_{\Omega(\bar{z})} Q_1(\bar{x}, \bar{x}^*, \bar{u}, \bar{z})\, P(\bar{z})\, d\Omega(\bar{z}), \tag{1.42}$$

où $\Omega(\bar{z})$ est le domaine de variation du vecteur de la perturbation $\bar{z}$, et $d\Omega(\bar{z})$ son élément infinitésimal.

Les systèmes optimaux aux critères d'optimalité statistiques ont fait l'objet d'étude de plusieurs ouvrages (cf. par exemple [1.10], [1.11], [1.12], [1.13]).

Les systèmes du troisième type, dits systèmes minimax, ne sont optimaux par rapport aux autres systèmes que dans le pire des cas. Autrement dit, le pire des résultats d'un système minimax est meilleur que le pire des résultats de tout autre système. Une telle formulation du problème convient parfois lorsque les distributions des probabilités a priori sont inconnues.

§ 4. Introduction de l'information sur l'objet commandé dans un organe de commande

Dans une branche importante de la classification des systèmes de commande on part du caractère de l'information sur l'objet gouverné B fournie à l'organe de commande A ou mise en mémoire par cet organe avant le départ du processus de commande.

Il faut d'abord distinguer les systèmes à information complète et ceux à information incomplète sur l'objet. D'après ce qui précède (cf. également fig. 1.1), l'information sur l'objet gouverné se compose de:

a) l'information sur son opérateur, c'est-à-dire sur la fonction $\bar{F}$ (1.6);

b) l'information sur la perturbation $\bar{z}$ subie par l'objet B;

c) l'information sur l'état de l'objet B, par exemple sur toutes les grandeurs $x_1, \ldots, x_n$ d'un objet dont le comportement est défini par les équations (1.7);

d) l'information sur le but de la commande, c'est-à-dire sur la fonctionnelle Q (1.21);

e) l'information sur la consigne $\bar{x}^*$.

L'information complète sur une relation quelconque équivaut à sa connaissance précise. Si toutes les composantes de l'information sur l'objet B indiquées ci-dessus sont connues d'avance ou fournies au cours du travail, le système considéré est dit système avec information complète sur l'objet. En pratique, quels que soient les systèmes de commande automatique, l'information sur l'objet ne peut pas être considérée comme complète; souvent l'absence de telle ou telle information est d'une grande importance. Revenons encore à la figure 1.1; on y voit plusieurs canaux par lesquels l'indétermination pénètre dans le système. Premièrement, c'est le canal de la consigne $\bar{x}^*$ qui dans de nombreux cas est inconnue d'avance. Un autre canal est celui de la perturbation $\bar{z}$ qui représente des variations aléatoires des caractéristiques de l'objet B, le plus souvent inconnues d'avance et ne se prêtant pas à la mesure directe. Il est fréquent que ce type d'indétermination est le plus important. Dans la majorité des cas c'est précisément à cause de l'indétermination des types considérés qu'on recourt à des systèmes de commande complexes. Si toute l'information sur l'objet était connue d'avance, on pourrait réaliser un système de commande en boucle ouverte où l'organe A serait muni d'un programme de commande établi d'avance. Aucun besoin n'est alors dans la théorie spéciale de la commande. Ajoutons aussi que dans un tel système hypothétique la boucle de réaction serait inutile.

Pourtant, une boucle de réaction, tout en étant un outil puissant pour accroître la résistance aux perturbations du système, présente en même temps un canal par lequel d'autres perturbations pénètrent dans le système. Par ce circuit on transmet à l'organe de commande A les données sur l'état de l'objet B, par exemple les coordonnées $x_1, \ldots, x_n$ de l'objet défini par les équations (1.17) ou les grandeurs équivalentes $x, \frac{dx}{dt}, \ldots, \frac{d^{n-1}x}{dt^{n-1}}$ si l'objet est caractérisé par l'équation (1.23). Souvent la grandeur x peut être mesurée avec une précision suffisante. Mais ce n'est pas du tout le cas de la dérivée première et surtout des dérivées d'ordre supérieur, ou, si leur mesure est possible, les données obtenues sont entachées d'erreurs non négligeables. La dérivation multiple de la fonction $x(t)$ est d'aucune aide. En effet, d'une part, tous les dérivateurs introduisent des erreurs. D'autre part, les faibles perturbations de hautes fréquences qui s'ajoutent inévitablement à la fonction $x(t)$ don-

nent lieu à des déformations des dérivées, d'autant plus grandes en général que l'ordre de la dérivée est plus élevé.

Il s'ensuit que les données sur l'état de l'objet sont fournies à l'organe de commande avec parfois de très grosses erreurs. Par conséquent, l'information sur l'état de l'objet est en pratique incomplète. On peut imaginer un schéma équivalent dans lequel les données sur l'état de l'objet seraient amenées à l'organe de commande après avoir passé par un canal aux perturbations aléatoires qui viendraient s'ajouter aux données. Un tel canal constitue précisément la troisième source de l'indétermination.

Il faut dire également que tout organe de commande réel A donne lieu à des perturbations aléatoires internes.

L'insuffisance de l'information sur l'objet rend nécessaire la collecte de l'information au cours de la régulation. En général, l'organe de commande d'un système automatique assure simultanément deux tâches étroitement liées mais différentes en principe. Premièrement, l'information qu'il reçoit lui permet de déterminer les caractéristiques et de juger de l'état de l'objet gouverné. Deuxièmement, une fois ces caractéristiques établies, il devient possible de définir ce qu'il faut entreprendre pour assurer la commande convenable. La première tâche consiste à étudier l'objet, la deuxième à l'amener à l'état imposé. Il se peut qu'un système simple n'assure qu'une de ces tâches, alors que l'autre ou bien n'est pas du tout accomplie ou bien est réduite au strict minimum. Dans les systèmes complexes les deux tâches sont obligatoires.

Le processus de l'étude de l'objet gouverné ressemble à celui de l'acquisition par l'homme de nouvelles connaissances. Ce peut être l'information communiquée par une autre personne, les résultats de l'observation ou les résultats d'une expérience. Les systèmes optimaux peuvent être également de trois types définis par le mode d'acquisition de l'information :

a) systèmes optimaux à information complète sur l'objet gouverné ou à information maximale au possible (cf. chapitres III et IV) ;

b) systèmes optimaux à information incomplète sur l'objet et son accumulation indépendante (ou passive) au cours du processus de commande (cf. chapitre V) ;

c) systèmes optimaux à information incomplète sur l'objet et son accumulation active pendant le processus de commande (commande duale) (cf. chapitre VI).

Nous avons admis dans ce qui suit que l'organe de commande A dispose d'une information complète a priori sur l'opérateur $\overline{F}$ de l'objet et sur le but de la commande, c'est-à-dire sur la fonctionnelle Q. Si, de plus, le système reçoit une information complète sur la consigne $\overline{x}^*$ (c'est-à-dire tous les renseignements possibles sur cette grandeur dans le passé, au présent et dans l'avenir), une

information complète sur la perturbation $\bar{z}$ (y compris la connaissance exacte de son comportement ultérieur) et, enfin, une information *courante* complète sur l'état $\bar{x}$ de l'objet à l'instant t considéré (et ceci permet, lorsque $\bar{u}(t)$ est donné, de prédire le comportement de l'objet dans l'avenir), on l'appelle alors système à information complète (dans l'organe de commande) sur l'objet gouverné.

Jusqu'à des temps récents la branche de la théorie des systèmes optimaux étudiant les systèmes à information complète évoluait indépendamment d'une autre branche apparue à peu près simultanément avec la première. Cette dernière partait non pas des actions de commande connues a priori, mais des caractéristiques statistiques des signaux d'entrée aléatoires. On ne disposait donc pas d'information complète. Le problème principal étudié par cette seconde branche intéresse le système dont le schéma fonctionnel est représenté sur la figure 1.4,*a*. D'abord le système entier est considéré comme un filtre F. La consigne x^* est injectée dans le filtre non pas directement, mais par un canal de liaison ou, en général, par un système donné H^* dans lequel elle s'ajoute à la perturbation aléatoire ou au bruit h^*. Ainsi le mélange y^* du signal avec le bruit est amené à l'entrée du filtre F. Le rôle du filtre est de fournir à sa sortie une grandeur x voisine au possible, dans un certain sens statistique, de x^* ou du résultat d'une certaine transformation de x^*.

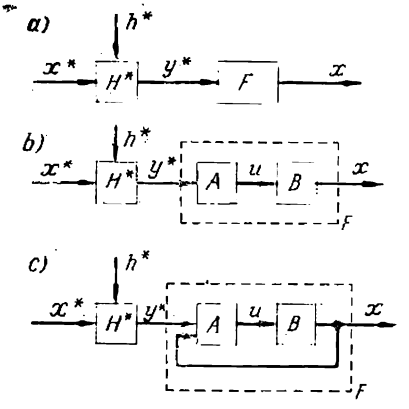


Fig. 1.4.

Après avoir résolu ce problème, on peut s'attaquer au problème suivant qui consiste à démembrer le filtre F en objet B et en organe de commande A. Ces parties de F peuvent être connectées en série comme sur la figure 1.4,*b* pour former un système en boucle ouverte, ou d'une autre façon quelconque. Ainsi, le filtre F de la figure 1.4,*c* comporte une boucle de réaction. Dans les cas courants l'objet B est donné d'avance, alors que l'algorithme de l'organe de commande A est à définir. S'il s'agit d'un filtre optimal linéaire, sa division en parties A et B n'entraîne pas de grosses difficultés *). Mais si le filtre optimal est non linéaire, le problème de dislocation se complique considérablement. Il est alors plus commode, avant de s'attaquer au calcul, d'imposer le schéma fonctionnel du filtre F pour lequel on recherche ensuite l'algorithme de la partie A.

*) Même dans ce cas le problème n'est pas toujours résolvable.

Toutefois l'étude des filtres non linéaires n'est qu'à ses premiers pas (cf. [1.10], [1.11], [1.17], [1.18]). Les ouvrages dans ce domaine, depuis les recherches de A. Kolmogorov [1.14] et N. Wiener [1.15], dans leur grande majorité posaient et résolvaient des problèmes relatifs aux systèmes optimaux linéaires (cf., par exemple, [1.10], [1.11], [1.12], [1.13], [1.16], [1.20]-[1.26]).

Le deuxième groupe de systèmes optimaux dont nous avons parlé plus haut se distingue par ce que le processus d'accumulation de l'information sur x^* est indépendant de l'algorithme ou, autrement dit, indépendant de la stratégie de l'organe de commande A. En effet, l'accumulation de l'information consiste à *observer* les valeurs y^* et à construire, d'après ces observations, des hypothèses sur le processus x^*. Le processus de l'observation ne dépend pas des décisions adoptées par l'organe A sur le caractère du processus x^*. On ne peut qu'utiliser correctement l'information fournie par l'observation, mais il est impossible de l'accroître, quelle que soit la stratégie de l'organe de commande. On appelle les systèmes de ce type systèmes optimaux à accumulation passive, ou indépendante, de l'information (indépendante de la stratégie de l'organe de commande). Si l'action x^* est irrégulière, ses caractéristiques peuvent être données de plusieurs façons:

a) $x^*(t)$ est une fonction d'une classe connue, par exemple,

$$x^*(t) = \sum_{i=1}^{\nu} C_i \psi_i(t), \tag{1.43}$$

où $\psi_i(t)$ sont les fonctions connues et C_i les variables aléatoires aux caractéristiques probabilistes connues. De plus, $h^*(t)$ est une fonction aléatoire dont les caractéristiques probabilistes sont également connues. Dans ce cas, plus la durée de l'observation de y^* à la sortie du canal H^* est grande, plus la prévision du comportement futur de $x^*(t)$ est précise, par exemple en précisant les valeurs des coefficients C_i dans la formule (1.43).

b) $x^*(t)$ est une fonction aléatoire; les caractéristiques probabilistes dont on dispose sont telles que la connaissance précise de son passé ne permet qu'une prévision de son futur. On ne peut donc qu'améliorer ces prévisions en observant $y^*(t)$ et en précisant d'après les résultats observés les valeurs anciennes et actuelles de $x^*(t)$. Pourtant, la prévision du comportement de la fonction $x^*(t)$ ne sera jamais aussi précise que l'on veut.

c) $x^*(t)$ est une fonction aléatoire à caractéristiques probabilistes totalement ou partiellement inconnues. Dans ces conditions le problème se ramène à établir ou à préciser par observation de la grandeur $y^*(t)$ les caractéristiques probabilistes de $x^*(t)$, ce qui rend possible de prévoir avec beaucoup plus de précision le comportement futur de $x^*(t)$ (cf., par exemple, [5.32, 5.33]).

Les ouvrages [6.1, 6.2] traitent certains problèmes du troisième groupe des systèmes optimaux. Ce groupe a des traits communs avec le premier et le deuxième groupe, tout en gardant sa particularité.

Le schéma fonctionnel donné par [6.1, 6.2] est reproduit sur la figure 1.5. L'action de commande $\bar{u}$ est amenée à l'objet B par le canal de liaison G dans lequel elle s'ajoute à la perturbation aléatoire (bruit) $\bar{g}$. On comprend qu'en général l'action $\bar{v}$ à l'entrée de l'objet B diffère de $\bar{u}$. L'information sur l'état $\bar{x}$ de l'objet passe ensuite

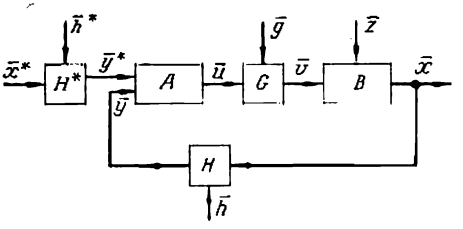


Fig. 1.5.

par le canal de liaison H où elle s'ajoute à la perturbation aléatoire (bruit) $\bar{h}$ pour se transformer en grandeur $\bar{y}$ et être amenée à l'entrée de l'organe de commande A. L'action extérieure $\bar{x}^*$ passe par le canal H^* de même que sur la figure 1.4. Le canal G avec la perturbation $\bar{g}$ peut être inclu dans l'objet B, $\bar{g}$ devenant alors composante du vecteur de la perturbation $\bar{z}$ de l'objet. Les autres blocs sont essentiellement indépendants. Dans le schéma à boucle fermée de la figure 1.5 se produisent les processus qui n'ont pas d'analogues dans les systèmes de commande à boucle ouverte. Il devient possible alors d'étudier la perturbation $\bar{z}$, c'est-à-dire, au fond, les caractéristiques de l'objet B à variation aléatoire, non par l'observation passive, mais par une méthode active, en faisant appel aux « expériences » rationnelles. C'est comme si l'on « tâtait » l'objet par des actions d'essai $\bar{u}$, alors que les résultats $\bar{y}$ de ces actions étaient analysés par l'organe de commande A. Le but de ces actions est de contribuer à une *étude* plus précise et plus rapide des caractéristiques de l'objet B et de favoriser ainsi l'établissement d'une meilleure loi de la commande de l'objet.

Pourtant l'action de commande est nécessaire non seulement pour étudier l'objet, mais encore pour l'*amener* à l'état imposé. C'est pourquoi dans le schéma de la figure 1.5 les actions de commande doivent avoir un caractère *dual*: elles doivent assurer à la fois l'étude et le guidage de l'objet, d'où le nom des systèmes de ce type — systèmes duals [cf. 6.1]. La dualité de la commande est le facteur physique principal qui distingue le troisième groupe

des systèmes optimaux des deux premiers. Pour le premier groupe, la commande duale est inutile puisque l'organe de commande reçoit une information complète sur l'objet. Pour le deuxième groupe la commande duale est impossible du fait que l'information est accumulée uniquement par l'observation et la vitesse du processus d'accumulation ne dépend aucunement de la stratégie de l'organe de commande.

La troisième branche de la classification des systèmes optimaux examinée dans ce paragraphe (cf. fig. 1.2) est intimement liée aux problèmes d'obtention de l'information nécessaire pour la commande. C'est précisément en ce point crucial que les méthodes statistiques viennent en aide à la théorie des systèmes optimaux.

§ 5. Formulation des problèmes de la théorie des systèmes optimaux

On voit de ce qui précède que la diversité kaléïdoscopique de types des systèmes optimaux ne fait pas obstacle à leur systématisation dans le cadre d'une classification relativement restreinte. Ceci permet une approche commune des problèmes que posent ces systèmes. Actuellement il devient possible d'édifier une théorie générale unique des systèmes optimaux, donnant la formulation des problèmes généraux et décrivant les méthodes de leur résolution. Dans ce qui suit nous formulons des problèmes généraux de cette théorie. L'étude des problèmes concrets, spécifiques à diverses branches de la théorie, fait l'objet des chapitres suivants.

Supposons donnés l'opérateur de l'objet $\overline{F}[\overline{u}, \overline{z}, t]$ et le but de la commande sous la forme de la fonctionnelle Q. Si les actions extérieures appliquées au système (fig. 1.5) sont aléatoires, supposons connues leurs caractéristiques probabilistes. Les fonctions régulières peuvent être aussi considérées du point de vue probabiliste, leur moyenne étant une fonction du temps, alors que la variance (mesure de la dispersion des valeurs d'une grandeur aléatoire d'une expérience à l'autre) est nulle. Soit dans le cas général les opérateurs des circuits H^*, H et G de la figure 1.5. Supposons également que les contraintes soient données sous la forme (1.16) ou (1.17), ou encore (1.19). Imposons-nous aussi la classe des fonctions $\overline{u}(t)$. Dans les cas courants, $\overline{u}(t)$ est considérée comme appartenant à la classe des fonctions continues par morceaux à nombre fini de points de discontinuité de première espèce dans un intervalle fini quelconque *).

Le problème consiste à trouver, pour des conditions données, l'algorithme ou, comme on dit parfois, la stratégie de l'organe de

*) Certains auteurs étudient les régimes dits « glissants » où ce nombre est infini [2.28, 2,29].

commande A, qui minimise le critère d'optimalité Q. La stratégie de ce type est dite optimale.

Dans le cas général une stratégie optimale peut être aléatoire. Cela signifie que l'organe de commande prend une décision aléatoire et peut fournir à sa sortie à l'instant t une grandeur aléatoire $\bar{u}$. Toutefois, la caractéristique probabiliste $\Gamma\{\bar{u}(t)\}$ de cette grandeur (sa densité de probabilité, par exemple) dépend d'une façon optimale de toute l'information antérieure reçue par l'organe de commande A et des actions qu'il a réalisées, c'est-à-dire de l'allure des fonctions $\bar{y}(\tau)$, $\bar{y}^*(\tau)$ et $\bar{u}(\tau)$ (fig. 1.5) pendant l'intervalle de temps τ depuis l'instant initial t_0 jusqu'à l'instant courant t, donc $t_0 \leqslant \tau < t$ (dans le cas particulier, $t_0 = -\infty$). La relation entre $\Gamma\{\bar{u}(t)\}$ et les fonctions indiquées ci-dessus ainsi que le temps t s'écrit sous la forme symbolique suivante:

$$\Gamma\{\bar{u}(t)\} = \Gamma\{\bar{u}(t) \mid \bar{y}^*(\tau),\ \bar{y}(\tau),\ \bar{u}(\tau),\ t\} \qquad (t_0 \leqslant \tau < t), \tag{1.44}$$

et se lit : fonction Γ de $\bar{u}(t)$ *sous la condition* que $\bar{y}^*$, $\bar{y}$ et $\bar{u}$ soient données dans l'intervale de t_0 à t.

La stratégie optimale peut être régulière dans ce cas particulier, l'une des valeurs possibles de $\bar{u}(t)$ a une probabilité égale à l'unité, alors que les probabilités de toutes les autres valeurs sont nulles. La stratégie optimale régulière s'exprime par la relation

$$\bar{u}(t) = K[\bar{y}^*(\tau),\ \bar{y}(\tau),\ \bar{u}(\tau),\ t] \quad (t_0 \leqslant \tau < t), \tag{1.45}$$

K étant également une fonctionnelle, c'est-à-dire un nombre défini par le type des fonctions $\bar{y}^*(\tau)$, $\bar{y}(\tau)$, $\bar{u}(\tau)$ dans l'intervalle $t_0 \leqslant \leqslant \tau < t$.

Dans le cas particulier des systèmes à information complète, lorsque les circuits H^*, H et G de la figure 1.5 sont absents, alors que $\bar{z}$ et $\bar{x}^*$ sont régulières et peuvent être inclues dans l'opérateur $\bar{F}$ de l'objet et le critère Q respectivement, l'expression générale (1.45) de l'algorithme d'un système optimal devient encore plus simple et se met sous la forme

$$\bar{u}(t) = \bar{K}[\bar{x}(\tau),\ \bar{u}(\tau),\ t] \quad (t_0 \leqslant \tau < t). \tag{1.46}$$

Si l'état courant de l'objet caractérisé par le vecteur $\bar{x}$ définit tout son comportement ultérieur indépendamment de la « préhistoire », c'est-à-dire des valeurs $\bar{x}(\tau)$ et $\bar{u}(\tau)$ pour $\tau < t$ (si, par exemple, l'objet est décrit par des équations du type (1.7) et ne contient ni retards ni relations non univoques), alors $\bar{u}(t)$ à l'instant t considéré est une fonction de la valeur de $\bar{x}$ à ce même instant t:

$$\bar{u}(t) = \bar{K}[\bar{x}(t),\ t]. \tag{1.47}$$

Enfin, si les équations de la trajectoire ne contiennent pas explicitement le temps t (les systèmes correspondants s'appellent souvent stationnaires), l'algorithme optimal se trouve sous la forme d'une fonction

$$\bar{u}(t) = \bar{K}\,[(\bar{x}(t)], \tag{1.48}$$

ou, sous une forme condensée,

$$\bar{u} = \bar{K}\,[\bar{x}]. \tag{1.49}$$

Dans le cas particulier d'une seule action de commande u, la formule de l'algorithme optimal s'écrit

$$u = K\,[\bar{x}]. \tag{1.50}$$

Il convient donc de trouver dans ce cas la fonction scalaire K de l'argument vectoriel $\bar{x}$, c'est-à-dire de la position du point représentatif (extrémité du vecteur $\bar{x}$) dans l'espace de phase du système. Autrement dit, K est la fonction de n variables $x_1, \ldots, x_n$.

Le problème de la recherche de la stratégie optimale s'appelle parfois problème de la synthèse du système optimal, ce qui n'est pas très juste, car il existe un très grand nombre de systèmes différents réalisant le même algorithme.

Le problème de la recherche de la stratégie optimale ou de l'algorithme d'un organe de commande est fondamental pour la théorie des systèmes optimaux ; depuis la fin des années 1940 il fait l'objet de différentes branches de cette théorie. Pourtant, il existe un autre problème, celui de la définition des processus optimaux, c'est-à-dire de la recherche des processus $\bar{u}(t)$ et $\bar{x}(t)$ en fonction du temps pour des conditions initiales $\bar{x}^0$. Ce n'est pas un problème principal, mais dans la plupart des cas il constitue une étape préalable dans la résolution du problème fondamental.

La recherche du processus optimal étant parfois intéressante par elle-même, nous allons en donner la formulation pour une classe définie de systèmes à information complète sur l'objet. Supposons que le comportement de l'objet soit décrit à l'aide de n équations de premier ordre du type (1.7) pour les coordonnées $x_1, \ldots, x_n$ ou à l'aide d'une équation vectorielle

$$\frac{d\bar{x}}{dt} = \bar{f}\,(\bar{x}, \bar{u}, t), \tag{1.51}$$

où f_i sont les fonctions continues et dérivables par rapport à leurs arguments.

Soit $\bar{x}^0$ la valeur initiale de $\bar{x}$ à l'instant $t = t_0$. La figure 1.6 représente le point aux coordonnées

$$\bar{x}^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \tag{1.52}$$

dans un espace de phase de dimension n d'un système aux coordonnées cartésiennes $x_1, \ldots, x_n$. Dans l'exemple schématisé par la figure 1.6 l'espace de phase est tridimensionnel. Mais tous les raisonnements qui suivent sont également valables pour un n quelconque.

L'action de commande $\bar{u}(t)$ produit un déplacement du système suivant une trajectoire dans l'espace de phase; à l'instant $t = T$ le point représentatif vient en $\bar{x}^{(T)}$. Soit $\bar{x}^{(T)}$ le point d'un certain sous-ensemble P des points de l'espace de phase. Le sous-ensemble P peut être, par exemple, à une dimension, c'est-à-dire représenter une courbe dans un espace de dimension n. Dans un cas particulier, le sous-ensemble P peut se confondre avec l'espace de phase. Le problème s'appelle alors problème à l'extrémité libre de trajectoire. Dans ce qui suit nous allons étudier deux cas : a) P dégénère en un point fixe $\bar{x}^{(T)}$, alors que la durée T n'est pas fixée d'avance; b) problème à extrémité libre de trajectoire lorsque T est fixé. Pourtant les méthodes décrites de résolution des problèmes sont applicables au cas général examiné dans le chapitre II.

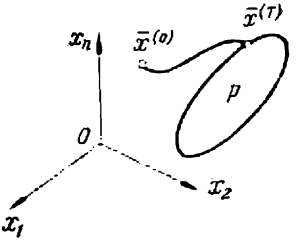


Fig. 1.6.

Supposons que les contraintes imposées au vecteur $\bar{u}$ soient du type (1.16) :

$$\bar{u} \in \Omega(\bar{u}). \tag{1.53}$$

Appelons action de commande *admissible* une fonction continue par morceaux $\bar{u}(t)$ vérifiant la contrainte (1.53). Le problème du processus optimal consiste à rechercher une action de commande admissible $\bar{u}(t)$ et une trajectoire respective $\bar{x}(t)$ de l'objet telles que la trajectoire du point représentatif $\bar{x}$ dans l'espace de phase qui passe de la position initiale $\bar{x}^{(0)}$ à la position $\bar{x}^{(T)}$, appartenant au sous-ensemble P, minimise une certaine fonctionnelle Q. Mettons cette dernière sous la forme d'une intégrale (avec $t_0 = 0$)

$$Q = \int_0^T G[\bar{x}(t), \bar{u}(t), t]\, dt, \tag{1.54}$$

où G est une fonction scalaire finie et généralement positive de x, $\bar{u}$ et t. Notons que formellement on peut faire disparaître t des expressions (1.51) et (1.54) en introduisant une coordonnée supplémentaire x_{n+1} avec la condition $(x_{n+1})_{t=0} = 0$ et

$$\frac{dx_{n+1}}{dt} = 1. \tag{1.55}$$

Puisque avec ces notations $x_{n+1} = t$, on peut remplacer partout t par x_{n+1}. Le nouveau système d'équations (1.51) et l'équation (1.55) ne comptent pas d'argument t sous une forme explicite, mais, par contre, sont caractérisés par $n + 1$ coordonnées $x_1, \ldots, x_n$, x_{n+1}. Ainsi on peut se borner à des équations ne contenant pas t explicitement, ce que nous allons faire souvent dans la suite.

Si dans le cas particulier du problème à point final fixé $\bar{x}^{(T)}$ et à la durée T non fixée d'avance on pose dans la formule (1.54) $G = 1$, on a $Q = T$. Cela signifie que la condition $Q = \min$ se transforme en $T = \min$. On se trouve alors devant le problème de *régulation en temps minimal* qui consiste à rechercher une loi de commande $\bar{u}(t)$ telle que le point représentatif $\bar{x}$ soit déplacé d'une position fixée $\bar{x}^{(0)}$ en une autre position fixée $\bar{x}^{(T)}$ en un temps minimal T.

Ajoutons aux coordonnées $x_1, \ldots, x_n$ la coordonnée x_0, avec $(x_0)_{t=0}$ et

$$\frac{dx_0}{dt} = G\,[x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ t]. \tag{1.56}$$

En comparant les expressions (1.54) et (1.56) on voit que

$$Q = (x_0)_{t=T} = x_0^{(T)}, \tag{1.57}$$

c'est-à-dire que le critère d'optimalité Q est égal à la valeur x_0 à l'instant final $t = T$. L'interprétation géométrique de ce fait (fig. 1.7) est donnée par un espace de dimension $(n + 1)$ aux coordonnées $x_0, x_1, \ldots, x_n$, dans lequel la trajectoire de phase passe du point $\bar{x}^0$ appartenant à l'hyperplan $(x_1, \ldots, x_n)$ jusqu'au point M dont les coordonnées $(x_1, \ldots, x_n)$ sont fixées, alors que la coordonnée x_0 ne l'est déjà plus. Par conséquent, M repose sur la perpendiculaire élevée du point $\bar{x}^{(T)}$ de l'hyperplan $x_0 = 0$ et parallèle à l'axe x_0. Il faut trouver une commande $\bar{u}(t)$ minimisant la valeur finale $x_0^{(T)}$ de la coordonnée x_0.

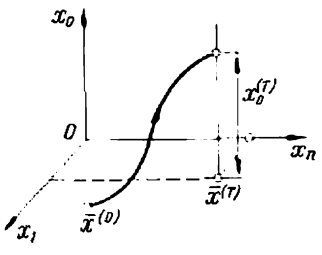


Fig. 1.7.

Dans les cas particuliers des problèmes de la théorie des systèmes optimaux à information incomplète, l'expression (1.45) de la fonctionnelle K peut être simplifiée de diverses façons. Supposons que la boucle de réaction soit absente. Alors $\bar{y}$ ne figure pas dans la formule (1.45) et

$$\bar{u}(t) = K\,[\bar{y}^*(\tau),\ \bar{u}(\tau),\ t] \qquad (t_0 \leqslant \tau < t). \tag{1.58}$$

Remarquant qu'aux instants passés $\bar{u}(\tau)$ était fonctionnelle d'un seul argument $\bar{y}^*(\tau)$, on peut écrire l'algorithme (1.58) comme

suit :

$$u(t) = K[y^*(\tau), t] \quad (t_0 \leqslant \tau < t). \tag{1.59}$$

Si le problème consiste à trouver l'algorithme du filtre F, représenté sur la figure 1.4,*a*, ou bien si l'on peut admettre que l'opérateur de l'objet B de la figure 1.4,*b* est unitaire (c'est-à-dire que la grandeur de sortie est égale à celle d'entrée), dans la formule (1.59) on peut remplacer $\bar{u}$ par $\bar{x}$. Si, de plus, le filtre F est stationnaire et $t_0 = -\infty$, il vient

$$\bar{x}(t) = K[\bar{y}^*(\tau)] \quad (-\infty < \tau < t). \tag{1.60}$$

Dans le cas particulier d'un filtre à canal unique, x et y^* sont des scalaires. Alors

$$x(t) = K[y^*(\tau)] \quad (-\infty < \tau < t). \tag{1.61}$$

Si, par exemple, le filtre F appartient à la classe des systèmes linéaires à coefficients constants, la formule pour K peut être mise sous la forme de l'intégrale de convolution

$$x(t) = \int_{-\infty}^{t} \varphi(t-\tau)\, y^*(\tau)\, d\tau, \tag{1.62}$$

où $\varphi(t)$ est la réponse impulsionnelle du filtre ou, comme on l'appelle souvent, la fonction de pondération. Soit $h^*(t)$ et $x^*(t)$ des processus aléatoires stationnaires (cf. chapitre II) et

$$y^*(t) = h^*(t) + x^*(t), \tag{1.63}$$

alors que le critère d'optimalité s'écrit

$$Q = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} [x^*(t) - x(t)]^2\, dt, \tag{1.64}$$

c'est-à-dire constitue une erreur quadratique moyenne. Le problème du système optimal prend alors la forme d'un problème qui consiste à définir la fonction de pondération $\varphi(t)$ d'un filtre physiquement réalisable vérifiant la condition

$$\varphi(t) \equiv 0 \quad (t < 0) \tag{1.65}$$

et minimisant le critère d'optimalité Q. Nous avons abouti ainsi au problème de filtration linéaire optimale résolu par N. Wiener dans [1.15]. Il se trouve que pour résoudre ce problème on n'a pas besoin de disposer de données exhaustives sur les caractéristiques probabilistes des processus $h^*(t)$ et $x^*(t)$. Il suffit de connaître les

fonctions de corrélation de ces processus, relativement simples à définir à base des expériences. Les ouvrages de A. Kolmogorov [1.14], A. Khintchine [1.19] et N. Wiener [1.15] ont jeté les fondements de la théorie de corrélation des systèmes optimaux linéaires. Nous ne nous attarderons pas à l'étude de cette théorie, car il existe déjà dans ce domaine un très grand nombre de monographies et de manuels. Les problèmes optimaux aux contraintes supplémentaires imposées à l'organe de commande A (dans le cas considéré, l'exigence de linéarité) ne sont pas examinés non plus dans notre ouvrage. Le lecteur s'intéressant aux divers aspects de cette théorie peut consulter [1.10]-[1.13], [1.16]-[1.18], [1.20]-[1.26].

CHAPITRE II

MÉTHODES MATHÉMATIQUES APPLIQUÉES À LA THÉORIE DES SYSTÈMES OPTIMAUX

§ 1. Eléments de calcul des probabilités

Dans le premier chapitre nous avons souligné le rôle des méthodes statistiques dans la théorie des systèmes optimaux. Pour la bonne compréhension de cette théorie il importe au préalable de traiter des éléments de base du calcul des probabilités et de la statistique mathématique. Ce paragraphe a pour objet de donner un résumé sommaire des concepts et des formules nécessités par l'exposé qui suit. Le lecteur désireux d'avoir des connaissances plus profondes de la théorie des probabilités et des méthodes statistiques peut s'adresser à plusieurs manuels et monographies, dont, entre autres, [2.1-2.4, 1.10-1.13]. De plus, certains éléments de statistique supplémentaires seront donnés dans les chapitres IV, V et VI.

La théorie des probabilités a pour objet l'étude de trois classes de phénomènes aléatoires. A la première on rapporte les phénomènes les plus simples dits *événements aléatoires*. La deuxième, plus complexe, comprend les *variables aléatoires*. Enfin la plus complexe traite des *processus aléatoires*.

Un événement aléatoire A est caractérisé par un nombre $p(A)$, dit probabilité, on sait de plus que $0 \leqslant p(A) \leqslant 1$. Pour un événement certain $p(A) = 1$, pour un événement impossible $p(A) = 0$.

Si les événements $A_1, A_2, \ldots, A_m$ sont incompatibles, c'est-à-dire si deux quelconques d'entre eux ne peuvent se produire au cours d'une même épreuve, la probabilité $p(A)$ de la réalisation au moins d'un seul de ces événements est exprimée par la formule

$$p(A) = \sum_{i=1}^{m} p(A_i). \tag{2.1}$$

L'événement A est souvent noté comme suit : $A = A_1 + \ldots + A_m$. Le signe « plus » remplace ici le terme logique « ou ».

Si l'un des événements A_i $(i = 1, \ldots, m)$ est certain, $p(A) = 1$ et

$$\sum_{i=1}^{m} p(A_i) = 1. \tag{2.2}$$

Les événements $A_1, \ldots, A_m$ dont les probabilités vérifient l'égalité (2.2) forment ce qu'on appelle *système complet.*

Supposons qu'au cours d'une série d'épreuves, effectuées dans des conditions déterminées, peut apparaître un événement aléatoire A avec une probabilité $p(A)$ que nous appellerons dès lors inconditionnelle, ou absolue.

Supposons également que ces épreuves font apparaître un autre événement aléatoire B avec une probabilité inconditionnelle $p(B)$. Prenons de toutes les N épreuves seulement les N_B qui ont donné lieu à l'événement B. Supposons que de toutes ces N_B épreuves seule la partie $N_{A|B}$ est caractérisée également par l'apparition de l'événement A. Le rapport $N_{A|B}/N_B$ s'appelle *fréquence* de l'événement A *sous la condition* de la réalisation de l'événement B, dite encore *fréquence conditionnelle* de l'événement A. Pour un grand nombre d'épreuves N_B le rapport $N_{A|B}/N_B$ diffère pratiquement peu d'un certain nombre que nous allons désigner par $p(A \mid B)$ et appeller *probabilité conditionnelle* de l'événement A (sous la condition de la réalisation de l'événement B).

Ce qui précède rend clair que $N_{A|B}$ est le nombre d'épreuves qui ont donné lieu aux deux événements A et B. Lorsque N est grand, le rapport $N_{A|B}/N$ ne diffère que peu de la probabilité de l'événement C consistant en une réalisation simultanée des événements A et B. On emploie souvent la notation $C = AB$, le signe de multiplication correspond ici à la conjonction « et » logique. Ainsi $p(C) = p(AB)$. Etant donné que

$$\frac{N_{A|B}}{N} = \frac{N_{A|B}}{N_B} \cdot \frac{N_B}{N}, \tag{2.3}$$

on peut écrire (par analogie) l'expression relative aux probabilités qui sont associées à ces fréquences:

$$p(AB) = p(A \mid B) \cdot p(B). \tag{2.4}$$

Les événements A et B étant équivalents, un raisonnement analogue conduit à une formule où A et B changent de place:

$$p(AB) = p(B \mid A) \cdot p(A). \tag{2.5}$$

Ainsi

$$p(AB) = p(A \mid B)\, p(B) = p(B \mid A)\, p(A). \tag{2.6}$$

Les événements A et B sont *indépendants* si l'apparition de l'un d'eux n'influe pas sur la probabilité de l'apparition de l'autre. L'événement A ne dépendant pas de B, $p(A \mid B) = p(A)$ et la formule (2.6) devient

$$p(AB) = p(A)\, p(B). \tag{2.7}$$

Les événements $A_1, \ldots, A_m$ sont dits *deux à deux indépendants* si deux quelconques d'entre eux sont indépendants. Mais si, de plus, les produits quelconques de ces événements $A_{i_1}, A_{i_2}, \ldots, A_{i_l}$ et

$A_{j_1}, A_{j_2}, \ldots, A_{j_q}$ sont également indépendants et ne comptent pas de facteurs communs, les événements $A_1, \ldots, A_m$ s'appellent *indépendants*. Il se peut, et nous insistons sur ce fait, que les concepts d'indépendance et d'indépendance deux à deux peuvent ne pas coïncider.

L'application successive de la formule (2.6) donne

$$p(A_1A_2 \ldots A_m) = p(A_1)\, p(A_2 \mid A_1) \ldots$$
$$\ldots p(A_m \mid A_1, A_2, \ldots, A_{m-1}). \qquad (2.8)$$

Dans le cas des événements indépendants l'expression (2.8) se simplifie et s'écrit

$$p(A_1A_2 \ldots A_m) = p(A_1)\, p(A_2) \ldots p(A_m). \qquad (2.9)$$

Soient $A_1, \ldots, A_m$ les événements incompatibles formant un système complet. Alors $A = A_1 + \ldots + A_m$ est un événement certain et la probabilité de n'importe quel autre événement B peut s'exprimer de la façon suivante :

$$p(B) = p(BA) = p[B(A_1 + \ldots + A_m)] =$$
$$= p(BA_1 + BA_2 + \ldots + BA_m) = \sum_{i=1}^{m} p(BA_i), \qquad (2.10)$$

les événements BA_i et BA_j étant incompatibles pour $i \neq j$. Par ailleurs, on tire de (2.5)

$$p(BA_i) = p(A_i)\, p(B \mid A_i). \qquad (2.11)$$

En portant cette expression dans (2.10) il vient

$$p(B) = \sum_{i=1}^{m} p(A_i)\, p(B \mid A_i). \qquad (2.12)$$

Par conséquent, (2.6) et (2.12) permettent de mettre l'expression de la probabilité conditionnelle $p(A_i \mid B)$ sous la forme suivante :

$$p(A_i \mid B) = \frac{p(A_i)\, p(B \mid A_i)}{p(B)} = \frac{p(A_i)\, p(B \mid A_i)}{\sum_{i=1}^{m} p(A_i)\, p(B \mid A_i)}. \qquad (2.13)$$

Cette formule établie en 1784 par Bayes porte son nom.

Une variable aléatoire est une notion plus complexe qu'un événement aléatoire. C'est une grandeur dont l'épreuve établit une et seulement une valeur tirée d'un ensemble de valeurs possibles. C'est pourquoi une variable aléatoire est définie non seulement par l'ensemble de ses valeurs possibles mais aussi par leurs probabilités. Soient $x_1, \ldots, x_n$ les valeurs possibles discrètes de la variable aléatoire ξ (n pouvant être fini ou infini). Il faut alors se donner les n probabilités du type $p_i = p(x_i)$, où p_i est la probabilité de l'événement aléatoire traduit par la réalisation de la valeur x_i

de la variable aléatoire ξ. Il est clair que

$$\sum_{i=1}^{n} p_i = 1 \tag{2.14}$$

du fait que, par définition de la variable aléatoire, pour des i différents, les événements $\xi = x_i$ sont incompatibles et, de plus, ils forment, toujours par définition, un système complet. Quand les valeurs de la variable aléatoire ξ sont quelconques dans un certain intervalle, sa caractéristique probabiliste doit être donnée sous une autre forme. La forme la plus générale est la *fonction de répartition* $F(x)$, dite encore *fonction cumulative*, donnant la probabilité de l'événement aléatoire $\xi < x$, où x est un seuil fixé à l'avance:

$$F(x) = p(\xi < x). \tag{2.15}$$

Connaissant $F(x)$ pour des x quelconques, il est facile de trouver a probabilité pour ξ de tomber dans l'intervalle $a \leqslant \xi < b$. En effet, puisque les événements $\xi < a$ et $a \leqslant \xi < b$ sont incompatibles et leur somme est l'événement $\xi < b$, on peut écrire

$$p(\xi < b) = p(\xi < a) + p(a \leqslant \xi < b). \tag{2.16}$$

On en déduit

$$p(a \leqslant \xi < b) = p(\xi < b) - p(\xi < a) = F(b) - F(a). \tag{2.17}$$

Si $F(x)$ est continue et dérivable dans tout l'intervalle $-\infty < x < \infty$, la variable aléatoire correspondante ξ est dite *continue*. Posons

$$P(x) = \frac{dF(x)}{dx} = F'(x). \tag{2.18}$$

La fonction $P(x)$ s'appelle *densité de probabilité* de la variable aléatoire ξ. Etant donné que

$$P(x) = \frac{dF(x)}{dx} = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = \\ = \lim_{\Delta x \to 0} \frac{p(x \leqslant \xi < x + \Delta x)}{\Delta x}. \tag{2.19}$$

$P(x)\,\Delta x$ est, à des infiniment petits d'ordre supérieur $o(\Delta x)$ près, la probabilité pour la variable aléatoire ξ de tomber dans l'intervalle infiniment petit $x \leqslant \xi < x + \Delta x$. Ensuite, la probabilité pour la variable ξ de se trouver dans l'intervalle $a \leqslant \xi < b$ est donnée par l'expression

$$p(a \leqslant \xi < b) = F(b) - F(a) = \int_a^b P(x)\,dx. \tag{2.20}$$

On en tire

$$\int_{-\infty}^{\infty} P(x)\,dx = F(\infty) - F(-\infty) = 1. \tag{2.21}$$

Des caractéristiques importantes bien que non exhaustives d'une variable aléatoire ξ sont ce qu'on appelle ses moments. Un moment d'ordre k est un nombre donné par l'intégrale

$$\alpha_k = \int_{-\infty}^{\infty} x^k P(x)\,dx. \tag{2.22}$$

Le moment d'ordre un α_1 dit *espérance mathématique* ou *valeur médiane* d'une variable aléatoire et noté m_ξ ou $M\{\xi\}$ présente un intérêt tout particulier :

$$m_\xi = M\{\xi\} = \alpha_1 = \int_{-\infty}^{\infty} xP(x)\,dx. \tag{2.23}$$

On appelle *moment centré* d'ordre k et on désigne par μ_k le moment d'ordre k de la différence $(\xi - m_\xi)$:

$$\mu_k = M\{(\xi - m_\xi)^k\} = \int_{-\infty}^{\infty} (x - m_\xi)^k P(x)\,dx. \tag{2.24}$$

Le moment centré d'ordre deux s'appelle *variance* ; il est noté D_ξ ou $D\{\xi\}$ et joue un rôle particulièrement important :

$$D_\xi = D\{\xi\} = M\{(\xi - m_\xi)^2\} = \int_{-\infty}^{\infty} (x - m_\xi)^2 P(x)\,dx. \tag{2.25}$$

La variance caractérise dans une certaine mesure la dispersion des valeurs d'une variable aléatoire ξ autour de sa valeur médiane m_ξ. La grandeur $\sqrt{D_\xi}$ se nomme écart-type et se note σ_ξ :

$$\sigma_\xi = \sqrt{D_\xi}. \tag{2.26}$$

Dans le cas général c'est une grandeur différente de la moyenne quadratique de ξ que nous allons désigner par $\xi_{\text{m. q}}$ et définir par la formule

$$(\xi_{\text{m. q.}})^2 = \alpha_2 = \int_{-\infty}^{\infty} x^2 P(x)\,dx. \tag{2.27}$$

Les expressions (2.25) et (2.27) ne coïncident que dans le cas où $m_\xi = 0$.

L'espérance mathématique d'une fonction quelconque $f(\xi) = \eta$ de la variable aléatoire ξ est donnée par l'expression

$$M\{\eta\} = M\{f(\xi)\} = \int_{-\infty}^{\infty} f(x) P(x)\, dx. \tag{2.28}$$

La loi de répartition des variables aléatoires la plus usitée est la *loi normale*, ou la *loi de Gauss-Laplace*, telle que

$$P(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{(x-m_x)^2}{2\sigma_x^2}\right\}, \tag{2.29}$$

m_x et σ_x sont ici des constantes égales respectivement à la valeur moyenne et à l'écart-type.

Pour la loi normale, la fonction de répartition est donnée par la formule

$$F(x) = \int_{-\infty}^{x} P(x)\, dx = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{(x-m_x)^2}{2\sigma_x^2}\right\} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m_x}{\sigma_x}} \exp\left\{-\frac{t^2}{2}\right\} dt. \tag{2.30}$$

La théorie des probabilités fait appel à la fonction

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{0}^{u} \exp\left\{-\frac{t^2}{2}\right\} dt. \tag{2.31}$$

Cette intégrale ne s'exprime pas par des fonctions élémentaires. La fonction $\Phi(u)$ est tabulée. Il est évident qu'elle est impaire, c'est-à-dire que $\Phi(-u) = -\Phi(u)$. La comparaison de (2.30) et de (2.31) montre que dans le cas de la loi normale

$$F(x) = \Phi\left(\frac{x-m_x)}{\sigma_x}\right). \tag{2.32}$$

Considérons un circuit non linéaire et sans inertie dont la relation entre la grandeur de sortie y et la grandeur d'entrée x s'écrit

$$y = f(x). \tag{2.33}$$

Si x est une variable aléatoire, y l'est également.

Supposons connue la densité de probabilité $P(x)$ d'une variable aléatoire x. Cherchons la densité de probabilité $P(y)$ de la variable aléatoire y. Nous employons ici le même symbole P pour la densité de probabilité de x et de y. Toutefois, la fonction $P(y)$ se distingue généralement de la fonction $P(x)$.

Proposons-nous tout d'abord de recherher la probabilité de l'appartenance de y à un certain domaine Ω_y : $y \in \Omega_y$. Ce domaine peut être, par exemple, l'intervalle $a \leqslant y < b$. Supposons que d'après (2.33) aux points y du domaine Ω_y correspondent les valeurs de x appartenant à un certain domaine Ω_x. La probabilité de l'événement $y \in \Omega_y$ est égale alors à celle de l'événement $x \in \Omega_x$. Par conséquent

$$p(y \in \Omega_y) = p(x \in \Omega_x) = \int_{\Omega_x} P(x)\, d\Omega_x, \tag{2.34}$$

où $d\Omega_x$ est un élément infiniment petit du domaine Ω_x et l'intégration porte sur tout le domaine Ω_x. Soit, par exemple, $P(x)$ suivant une loi normale de moyenne m_x nulle

$$P(x) = P_0(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma_x^2}\right\}. \tag{2.35}$$

Soit ensuite

$$y = A + Bx, \tag{2.36}$$

où A et $B > 0$ sont des constantes.

Cherchons le domaine Ω_x correspondant à l'intervalle infiniment petit des valeurs de y entre y_1 et $y_1 + dy$. Le domaine Ω_x est défini évidemment par la condition $x_1 \leqslant x < x_1 + dx$, où

$$x_1 = \frac{y_1 - A}{B} \tag{2.37}$$

et

$$dx = \frac{dy}{B}. \tag{2.38}$$

La formule (2.34) entraîne

$$p(y_1 \leqslant y < y_1 + dy) = P(y_1)\, dy = p(x_1 \leqslant x < x_1 + dx) =$$
$$= P_0(x_1)\, dx = P_0\left(\frac{y_1 - A}{B}\right) \frac{dy}{B}. \tag{2.39}$$

Si l'on remplace y_1 par y, il vient

$$P(y) = \frac{1}{B} P_0\left(\frac{y-A}{B}\right) = \frac{1}{B} \cdot \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{\left(\frac{y-A}{B}\right)^2}{2\sigma_x^2}\right\} =$$
$$= \frac{1}{B\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{(y-A)^2}{2(B\sigma_x)^2}\right\}. \tag{2.40}$$

Posons

$$\left.\begin{aligned} \sigma_y &= B\sigma_x, \\ m_y &= A. \end{aligned}\right\} \tag{2.41}$$

La formule (2.40) se met alors sous la forme

$$P(y) = \frac{1}{\sigma_y \sqrt{2\pi}} \exp\left\{-\frac{(y-m_y)^2}{2\sigma_y^2}\right\}. \tag{2.42}$$

Cette expression correspond exactement à la formule (2.29). On en déduit que y possède une loi normale de répartition de moyenne A et d'écart-type $B\sigma_x$. Plus B est petit, plus la courbe $P(y)$ est concentrée autour de la valeur moyenne de $y = A$.

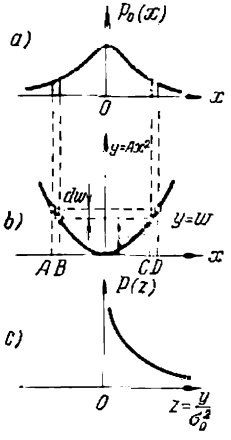


Fig. 2.1.

La caractéristique d'un circuit non linéaire peut être non univoque. Il se peut qu'à une même valeur de y correspondent plusieurs valeurs de x. Soit

$$y = Ax^2, \tag{2.43}$$

où $A > 0$ (fig. 2.1,*a*). Soit x la variable aléatoire normale amenée à l'entrée du circuit dont la loi $P_0(x)$ est donnée par la formule (2.35). Cherchons la probabilité pour y de se trouver entre les valeurs $w > 0$ et $w + dw$, c'est-à-dire la grandeur

$$p(w \leqslant y < w + dw) = P(w)\, dw, \tag{2.44}$$

où P est la densité de probabilité de y. Les figures 2.1,*a* et *b* montrent que cette probabilité est la somme des probabilités pour x de se trouver dans les intervalles AB et CD :

$$\begin{aligned} P(w)\,dw = p(w < y < w + dw) &= \\ = p\left(+\sqrt{\frac{w}{A}} < x < +\sqrt{\frac{w+dw}{A}}\right) &+ \\ + p\left(-\sqrt{\frac{w+dw}{A}} < x < -\sqrt{\frac{w}{A}}\right) &= \\ = 2p\left(+\sqrt{\frac{w}{A}} < x < +\sqrt{\frac{w+dw}{A}}\right). & \end{aligned} \tag{2.45}$$

Cette dernière transformation est légitime du fait que la répartition $P_0(x)$ est symétrique. Par conséquent les deux surfaces infiniment petites hachurées de la figure 2.1,*a* sont égales entre elles. Etant donné que

$$\sqrt{\frac{w+dw}{A}} \approx \sqrt{\frac{w}{A}} + \frac{dw}{2\sqrt{Aw}} \tag{2.46}$$

aux infiniment petits d'ordre supérieur près, (2.45) peut s'écrire

$$P(w)\,dw = 2P_0\left(x = +\sqrt{\frac{w}{A}}\right) \cdot \frac{dw}{2\sqrt{w}} \cdot \frac{1}{\sqrt{A}} \quad (w > 0), \tag{2.47}$$

ce qui entraîne

$$P(w) = \frac{1}{\sqrt{Aw}} P_0\left(+\sqrt{\frac{w}{A}}\right) \quad (w > 0). \tag{2.48}$$

Remplaçons w par y et exprimons $P_0(x)$ d'après (2.35), il vient

$$P(y) = \frac{1}{\sqrt{Ay}} \cdot \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left\{-\frac{y}{2\sigma_x^2 A}\right\} =$$

$$= \frac{1}{\sigma_0 \sqrt{2\pi y}} \exp\left\{-\frac{y}{2\sigma_0^2}\right\}, \tag{2.49}$$

où

$$\sigma_0 = \sigma_x \sqrt{A}. \tag{2.50}$$

La formule (2.49) n'est applicable que pour $y > 0$. Les valeurs $y < 0$ sont impossibles (fig. 2.1,*b*), alors pour $y < 0$ $P(y) = 0$. Sur la figure 2.1,*c* cette courbe est représentée en fonction de la grandeur $z = \frac{y}{\sigma_0^2}$. Il résulte de (2.49) que quand $z > 0$

$$P(z) = P\left(\frac{y}{\sigma_0^2}\right) = \frac{1}{\sigma_0^2 \sqrt{2\pi z}} \exp\left\{-\frac{z}{2}\right\}. \tag{2.51}$$

Cette loi diffère nettement de la loi normale.

Pour un ensemble de variables aléatoires $\xi_1, \ldots, \xi_n$ il est possible de donner aussi bien la fonction de répartition que la densité de probabilité.

L'ensemble des variables aléatoires $\xi_1, \ldots, \xi_n$ peut être envisagé comme les coordonnées cartésiennes d'un point ou comme les composantes d'un vecteur aléatoire $\overline{\xi} = (\xi_1, \ldots, \xi_n)$ dans un espace à n dimensions. La probabilité $p(A)$ de l'événement A, qui consiste à retrouver l'extrémité du vecteur $\overline{\xi}$ dans le domaine Ω_A de cet espace, est évidemment égale à la somme des probabilités pour l'extrémité considérée de joindre les volumes infiniment petits $d\Omega_A = dx_1 \ldots dx_n$ de ce domaine:

$$p(A) = \int_{\Omega_A} P(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = \int_{\Omega_A} P(\overline{x})\, d\Omega_A. \tag{2.52}$$

Nous avons introduit ici une notation abrégée $P(\overline{x})$ pour $P(x_1, \ldots, x_n)$.

La probabilité pour l'extrémité du vecteur $\overline{\xi}$ de rejoindre un point quelconque de l'espace à n dimensions étant égale à l'unité, il vient

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} P(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = 1. \tag{2.53}$$

Pour obtenir la densité de probabilité $P_i(x_i)$ d'une variable ξ_i il faut intégrer $P(x_1, \ldots, x_n)$ dans tout le domaine de variation des autres variables. Par exemple,

$$P_1(x_1) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} P(x_1, \ldots, x_n)\, dx_2 \ldots dx_n. \qquad (2.54)$$

Si les variables $\xi_1, \ldots, \xi_n$ sont indépendantes, d'après (2.9), la probabilité totale $p(A)$ de l'événement A envisagé ci-dessus est égale au produit des probabilités prises isolément pour les variables aléatoires ξ_i $(i = 1, \ldots, n)$. Par suite dans le cas considéré

$$P(x_1, \ldots, x_n) = P_1(x_1) \cdot P_2(x_2), \ldots, P_n(x_n). \qquad (2.55)$$

Les densités de probabilité P_i des différentes ξ_i peuvent être différentes.

Si les variables aléatoires ξ et η sont dépendantes, la fixation de l'une d'elles influe sur la répartition des probabilités de l'autre. Soit $P(y \mid x)\, dy$ la probabilité de la variable aléatoire η de tomber dans l'intervalle donné $y, y + dy$ à condition que la variable aléatoire ξ a une certaine valeur fixe x. Appelons *densité de probabilité conditionnelle* la densité de probabilité $P(y \mid x)$ pour η avec $\xi = x$. La probabilité pour η d'avoir une valeur quelconque étant égale à 1, on a

$$\int_{-\infty}^{\infty} P(y \mid x)\, dy = 1. \qquad (2.56)$$

Connaissant la densité de probabilité conjointe $P(x, y)$ des variables ξ et η, il est facile de trouver la densité conditionnelle $P(y \mid x)$. En effet, le théorème du produit des probabilités permet d'écrire

$$P(x, y)\, dx\, dy = [P(x)\, dx] \cdot [P(y \mid x)\, dy], \qquad (2.57)$$

où $P(x)$ est la densité de probabilité inconditionnelle de ξ. On en déduit

$$P(y \mid x) = \frac{P(x, y)}{P(x)}. \qquad (2.58)$$

Les formules (2.56) et (2.58) se généralisent aisément au cas des grandeurs vectorielles $\overline{\xi}$ et $\overline{\eta}$, c'est-à-dire des ensembles $\xi_1, \ldots, \xi_n$ et $\eta_1, \ldots, \eta_m$.

Pour l'ensemble des variables aléatoires $\xi_1, \ldots, \xi_n$ on peut également introduire la notion de valeur moyenne. Par exemple, la valeur moyenne de ξ_k peut être calculée d'après la formule

$$M\{\xi_k\} = m_{\xi_k} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} x_k P(x_1, \ldots, x_n)\, dx_1 \ldots dx_n =$$

$$= \int_{\Omega(\overline{x})} x_k P(\overline{x})\, d\Omega(\overline{x}), \qquad (2.59)$$

où $\Omega(\bar{x})$ est tout l'espace de dimension n des points des extrémités du vecteur $\bar{\xi}$, alors que $d\Omega(\bar{x}) = dx_1 \ldots dx_n$ est un élément infiniment petit de cet espace.

Pour l'ensemble des variables $\xi_1, \ldots, \xi_n$ les moments centrés d'ordre deux sont déterminés par la formule

$$M_{jk}\{\xi_j\xi_k\} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} [x_j - M\{\xi_j\}] \times$$
$$\times [x_k - M\{\xi_k\}]\, P(x_1, \ldots, x_n)\, dx_1 \ldots dx_n. \quad (2.60)$$

Pour $j = k$ cette formule donne la variance d'une variable aléatoire ξ_j et pour $j \neq k$ la covariance des variables aléatoires ξ_k et ξ_j. Quand ces variables sont indépendantes, il est facile de montrer que la covariance est nulle (en général, l'inverse n'est pas vrai : une covariance nulle ne permet pas d'affirmer que les variables ξ_j et ξ_k sont indépendantes). Le rapport sans dimension

$$R_{jk} = \frac{M_{jk}}{\sqrt{M_{jj}M_{kk}}} \quad (2.61)$$

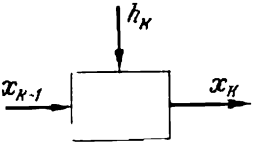


Fig. 2.2.

s'appelle *coefficient de corrélation* des variables aléatoires ξ_j et ξ_k.

Les notions de moyennes conditionnelles s'établissent d'une façon analogue. Ainsi, la moyenne conditionnelle $M\{\eta \mid \xi\}$ de la variable aléatoire η pour $\xi = x$, connaissant la densité conditionnelle $P(y \mid x)$ de η pour ξ fixée, s'obtient d'après la formule générale du type (2.58) :

$$M\{\eta \mid \xi\} = M_{\eta|\xi} = \int_{-\infty}^{\infty} yP(y \mid x)\, dy = \frac{1}{P(x)} \int_{-\infty}^{\infty} yP(x, y)\, dy. \quad (2.62)$$

Considérons un circuit sans inertie d'un système automatique (fig. 2.2) dont la grandeur de sortie x_k est fonction de deux grandeurs d'entrée h_k et x_{k-1} :

$$x_k = f_k(x_{k-1}, h_k). \quad (2.63)$$

Soit par exemple x_{k-1} une grandeur d'entrée du circuit et h_k une perturbation.

Admettons que x_{k-1} soit fixée et h_k soit une variable aléatoire de densité de probabilité $P(h_k)$. D'après la formule (2.63) on peut obtenir alors la densité de probabilité conditionnelle $P(x_k \mid x_{k-1})$ pour x_k à condition de donner une certaine valeur fixe à x_{k-1}, considérée dans le cas concerné comme un paramètre.

Maintenant la moyenne conditionnelle de x_k, x_{k-1} étant fixe, peut s'écrire

$$M\{x_k \mid x_{k-1}\} = M_{x_k|x_{k-1}} = \int_{-\infty}^{\infty} x_k P(x_k \mid x_{k-1})\, dx_k. \qquad (2.64)$$

Dans le cas de plusieurs éléments de ce type couplés en série (fig. 2.3), la moyenne conditionnelle du dernier élément est donnée par la formule

$$M\{x_n \mid x_{n-1}\} = M_{x_n|x_{n-1}} = \int_{-\infty}^{\infty} x_n P(x_n \mid x_{n-1})\, dx_n. \qquad (2.65)$$

Pourtant, même avec une grandeur d'entrée x_{n-2} fixe, la perturbation aléatoire h_{n-1} rend aléatoire la grandeur x_{n-1}. Si l'on connaît les propriétés de l'avant-dernier élément, on peut calculer

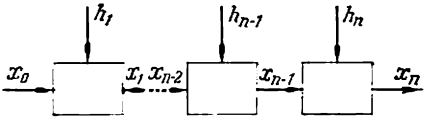


Fig. 2.3.

la densité conditionnelle $P(x_{n-1} \mid x_{n-2})$. Par économie des notations, nous avons employé ici la même lettre P que pour la fonction $P(x_n \mid x_{n-1})$ bien que généralement toutes ces fonctions puissent être différentes. Si x_{n-1} est une variable aléatoire, la moyenne (2.65) doit être considérée également comme une variable aléatoire. La valeur moyenne de cette dernière s'écrit à son tour

$$M\{x_n \mid x_{n-2}\} = M_{x_n|x_{n-2}} = M_{x_{n-1}|x_{n-2}}\{M_{x_n|x_{n-1}}\} =$$

$$= \int_{-\infty}^{\infty} P(x_{n-1} \mid x_{n-2}) \left[\int_{-\infty}^{\infty} x_n P(x_n \mid x_{n-1})\, dx_n \right] dx_{n-1} =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_n P(x_n \mid x_{n-1})\, P(x_{n-1} \mid x_{n-2})\, dx_n\, dx_{n-1} =$$

$$= \int_{\Omega(x_n,\, x_{n-1})} x_n P(x_n \mid x_{n-1})\, P(x_{n-1} \mid x_{n-2})\, d\Omega(x_n, x_{n-1}), \qquad (2.66)$$

où $\Omega(x_n, x_{n-1})$ est la notation de tout le domaine de variation des variables x_n et x_{n-1}, alors que $d\Omega(x_n, x_{n-1}) = dx_n\, dx_{n-1}$ désigne un de ses éléments infiniment petits. En raisonnant par récurrence,

tout en remontant la chaîne des éléments de la figure 2.3 de la sortie à l'entrée, nous arrivons finalement à la formule

$$M\{x_n \mid x_0\} = M_{x_n|x_0} = \int\limits_{\Omega(x_1, x_2, \ldots, x_n)} x_n P(x_n \mid x_{n-1}) \times$$
$$\times P(x_{n-1} \mid x_{n-2}) \ldots P(x_1 \mid x_0)\, d\Omega(x_1, \ldots, x_n) =$$
$$= \int\limits_{\Omega(\bar{x})} x_n \left[\prod_{k=1}^{n} P(x_k \mid x_{k-1})\right] d\Omega(\bar{x}), \quad (2.67)$$

où $\Omega(\bar{x})$ est le domaine de variation du vecteur $\bar{x} = (x_1, \ldots, x_n)$ et $d\Omega(\bar{x})$ son élément infiniment petit.

Passons à l'étude des processus aléatoires qui sont les événements aléatoires les plus complexes. Un processus aléatoire dit également *probabiliste* ou *stochastique* est une fonction aléatoire du temps, c'est-à-dire une fonction telle qu'à tout instant elle est une variable aléatoire. On peut aussi définir un processus aléatoire comme un ensemble de variables aléatoires $\xi(t)$, fonction de l'argument réel t.

Les observations isolées d'un processus aléatoire $\xi(t)$ se déroulant dans des systèmes du même type, c'est-à-dire dans des conditions contrôlées invariables de l'épreuve, donnent chaque fois des fonctions $x(t)$ différentes, des *échantillons* ou *réalisations* différents du processus aléatoire. La plus simple des caractéristiques probabilistes de ce processus est la loi de répartition unidimensionnelle $P_1(x_1, t_1)$, c'est-à-dire la densité de probabilité de la valeur $\xi(t_1)$ du processus à l'instant $t = t_1$. L'expression $P_1(x_1, t_1)\, dx_1$ traduit la probabilité de l'événement $x_1 < \xi(t_1) < x_1 + dx_1$. La fonction plus complexe $P_2(x_1, t_1; x_2, t_2)$ est la densité de probabilité bidimensionnelle d'une répartition conjointe de deux variables aléatoires : de la valeur $\xi(t_1)$ du processus à l'instant $t = t_1$ et de la valeur $\xi(t_2)$ à l'instant $t = t_2$. L'expression $P_2(x_1, t_1; x_2, t_2)\, dx_1\, dx_2$ est la probabilité de la réalisation des conditions $x_1 < \xi(t_1) < x_1 + dx_1$ et $x_2 < \xi(t_2) < x_2 + dx_2$. En général, toutes les n valeurs $\xi(t_1)$, $\xi(t_2)$, ..., $\xi(t_n)$ d'un processus aléatoire aux instants $t_1, \ldots, t_n$ peuvent être considérées comme n variables aléatoires. L'ensemble de ces dernières est caractérisé par la densité de probabilité conjointe $P_n[x_1, t_1; x_2, t_2; \ldots; x_n, t_n]$, et $P_n\, dx_1 \ldots dx_n$ est la probabilité de la réalisation des conditions

$$x_1 < \xi(t_1) < x_1 + dx_1, \ldots, x_n < \xi(t_n) < x_n + dx_n.$$

Les caractéristiques probabilistes d'un processus aléatoire sont entièrement connues si l'on connaît la fonction P_n pour n'importe quel n.

Le type le plus simple des processus aléatoires est celui où les valeurs de $\xi(t)$ à des instants différents sont indépendantes. Pour un

tel processus on a

$$P_n [x_1, t_1; x_2, t_2; \ldots; x_n, t_n] = \\ = P_1 (x_1, t_1) \cdot P_1 (x_2, t_2) \ldots P_1 (x_n, t_n). \qquad (2.68)$$

P_1 est ici une loi de distribution unidimensionnelle à partir de laquelle on peut construire à l'aide de la formule (2.68) une loi à n dimensions.

Un autre exemple est fourni par le processus aléatoire markovien (du nom de A. Markov, mathématicien éminent, qui a été le premier à étudier les processus de ce type). Pour un processus markovien, toute densité de probabilité P_n s'obtient à partir de $P_2 (x_1, t_1; x_2, t_2)$. Soit $P (x_n, t_n \mid x_{n-1}, t_{n-1})$ la densité de probabilité conditionnelle de $\xi (t_n)$ à l'instant $t = t_n$ sous la condition qu'à l'instant t_{n-1} précédant t_n $(t_{n-1} < t_n)$, on eût $\xi (t_{n-1})$ égale à x_{n-1}. Dans le cas d'un processus markovien cette densité de probabilité conditionnelle reste inchangée même si l'on connaît de plus les valeurs $x_{n-2}, \ldots, x_1$ du processus $\xi (t)$ aux instants précédents $t_{n-2}, \ldots, t_1$ $(t_n > t_{n-1} > t_{n-2} > \ldots > t_1)$. Fixons l'instant $t = t_{n-1}$ et rapportons au futur tous les $t > t_{n-1}$ et au passé tous les $t < t_{n-1}$. On peut alors considérer que les caractéristiques probabilistes d'un processus markovien évaluées pour les instants futurs sont définies par la valeur x_{n-1} du processus à l'instant donné t_{n-1} et ne dépendent pas de la « préhistoire » de ce processus, c'est-à-dire de ses valeurs antérieures lorsque t était inférieur à t_{n-1}.

Montrons que dans ce cas tout P_n peut être exprimé par P_2. En premier lieu la densité de probabilité unidimensionnelle P_1 s'obtient à partir de P_2 à l'aide de la formule (2.54) :

$$P_1 (x_1, t_1) = \int_{x_2=-\infty}^{x_2=\infty} P_2 (x_1, t_1; x_2, t_2)\, dx_2. \qquad (2.69)$$

Généralement, toute densité P_i peut être obtenue à partir d'une densité P_k $(k > i)$ par intégration par rapport aux variables qui ne figurent pas dans P_i.

Vu l'indépendance de la densité de probabilité conditionnelle de la « préhistoire » du processus, on peut écrire en posant $n > 2$:

$$P (x_n, t_n) \mid x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = P (x_n, t_n \mid x_{n-1}, t_{n-1}). \qquad (2.70)$$

Nous avons employé ici la même lettre P pour désigner deux densités de probabilité différentes des premier et second membres de l'expression (2.70) qui, pour un processus markovien, sont pourtant égales.

En vertu du théorème de la probabilité d'un événement composé (2.6) on a l'égalité

$$P_2 (x_n, t_n; x_{n-1}, t_{n-1}) = P (x_n, t_n \mid x_{n-1}, t_{n-1}) \cdot P_1 (x_{n-1}, t_{n-1}). \qquad (2.71)$$

La formule (2.70) peut donc s'écrire

$$P(x_n, t_n \mid x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = \frac{P_2(x_n, t_n; x_{n-1}, t_{n-1})}{P_1(x_{n-1}; t_{n-1})}. \quad (2.72)$$

Mais le théorème de la probabilité d'un événement composé entraîne

$$P_n(x_1, t_1; x_2, t_2; \ldots; x_n, t_n) =$$
$$= P_{n-1}(x_1, t_1; \ldots; x_{n-1}, t_{n-1}) P(x_n, t_n \mid x_{n-1}, t_{n-1}; \ldots; x_1, t_1) =$$
$$= P_{n-1}(x_1, t_1; \ldots; x_{n-1}, t_{n-1}) \cdot \frac{P_2(x_n, t_n; x_{n-1}, t_{n-1})}{P_1(x_{n-1}, t_{n-1})}. \quad (2.73)$$

On voit que dans le cas concerné, P_1 et P_2 étant connues, une fonction P_n quelconque se ramène à P_{n-1} qui à son tour s'exprime à l'aide de P_{n-2}, etc., en aboutissant finalement à la fonction P_2. Dans le cas d'un processus markovien, une fonction quelconque P_n s'exprime donc par l'intermédiaire de P_2.

On distingue des processus aléatoires stationnaires et non stationnaires. Un processus aléatoire est dit stationnaire si un décalage de tous les points $t_1, t_2, \ldots, t_n$ le long de l'axe du temps d'une même grandeur t_0 ne change pas les densités de probabilité respectives P_n. Un processus stationnaire vérifie l'égalité

$$P_n(x_1, t_1; x_2, t_2; \ldots; x_n, t_n) =$$
$$= P_n(x_1, t_1 + t_0; x_2, t_2 + t_0; \ldots; x_n, t_n + t_0). \quad (2.74)$$

Donc le caractère statistique d'un processus aléatoire stationnaire reste inchangé dans le temps. Un processus aléatoire stable est analogue au processus stable. Pour la densité de probabilité unidimensionnelle P_1 la formule (2.74) donne

$$P_1(x_1, t_1) = P_1(x_1, t_1 + t_0). \quad (2.75)$$

On voit que P_1 ne dépend nullement de t, c'est-à-dire que

$$P_1(x_1, t_1) = P_1(x_1). \quad (2.76)$$

Pour une densité de probabilité bidimensionnelle l'égalité (2.74) s'écrit

$$P_2(x_1, t_1; x_2, t_2) = P_2(x_1, t_1 + t_0; x_2, t_2 + t_0). \quad (2.77)$$

Cette relation montre que P_2 dépend non pas des instants t_1 et t_2, mais de leur différence $t_2 - t_1 = \tau$:

$$P_2(x_1, t_1; x_2, t_2) = P_2(x_1, x_2, \tau). \quad (2.78)$$

Deux processus aléatoires $\xi(t)$ et $\eta(t)$ liés sont caractérisés par des densités de probabilité conjointes des variables $\xi(t_i)$ et $\eta(t_j)$. Par exemple, $P_2(x_1, t_1; y_2, t_2)\, dx_1\, dy_2$ est la probabilité du fait que $x_1 < \xi(t_1) < x_1 + dx_1$ et $y_2 < \eta(t_2) < y_2 + dy_2$.

Si toutes les fonctions P_n ne dépendent pas du décalage des instants $t_1, t_2, \ldots$ d'une même valeur t_0, les processus $\xi(t)$ et $\eta(t)$ sont dits *processus stationnairement liés*.

La valeur moyenne ou l'espérance mathématique d'un processus aléatoire $\xi(t)$ à l'instant $t = t_1$ est définie comme l'espérance mathématique de la variable aléatoire $\xi(t_1)$. Pour simplifier l'écriture on peut remplacer $\xi(t_1)$ par $x(t_1)$. Il vient alors

$$M\{x(t_1)\} = \int_{-\infty}^{\infty} x_1 P_1(x_1, t_1)\, dx_1. \tag{2.79}$$

D'une façon analogue on définit le carré moyen

$$M\{x^2(t_1)\} = \int_{-\infty}^{\infty} x_1^2 P_1(x_1, t_1)\, dx_1. \tag{2.80}$$

Pour un processus aléatoire stationnaire la condition (2.76) entraîne

$$M\{x(t_1)\} = \int_{-\infty}^{\infty} x_1 P_1(x_1)\, dx_1 \tag{2.81}$$

et

$$M\{x^2(t_1)\} = \int_{-\infty}^{\infty} x_1^2 P_1(x_1)\, dx_1. \tag{2.82}$$

La moyenne du produit des variables $x(t_1) = x_1$ et $x(t_2) = x_2$, désignée par $K_x(t_1, t_2)$, s'appelle fonction *de corrélation ou d'autocorrélation.* Par définition de la moyenne

$$K_x(t_1, t_2) = M\{x(t_1)\, x(t_2)\} = \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 P_2(x_1, t_1;\, x_2, t_2)\, dx_1\, dx_2. \tag{2.83}$$

Pour un processus aléatoire stationnaire la condition (2.78) permet de simplifier la formule (2.83) car K_x ne dépend que de $\tau = t_2 - t_1$:

$$K_x(\tau) = M\{x(t_1)\, x(t_1 + \tau)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 P_2(x_1, x_2, \tau)\, dx_1\, dx_2. \tag{2.84}$$

Dans le cas particulier pour $\tau = 0$ la formule (2.84) donne, compte tenu de (2.82),

$$K_x(0) = [K_x(\tau)]_{\tau=0} = M\{x^2(t_1)\}. \tag{2.85}$$

D'une manière analogue à (2.83) on trouve la fonction de corrélation mutuelle des processus $x(t)$ et $y(t)$:

$$K_{xy}(t_1, t_2) = M\{x(t_1)\, y(t_2)\} = \\ = \int_{\infty}^{\infty} \int_{-\infty}^{\infty} x_1 y_2 P_2(x_1, t_1; y_2, t_2)\, dx_1\, dy_2. \quad (2.86)$$

Si les processus $x(t)$ et $y(t)$ sont stationnaires et stationnairement liés, on a

$$P_2(x_1, t_1; y_2, t_2) = P_2(x_1, y_2, \tau), \quad (2.87)$$

où $\tau = t_2 - t_1$. Alors K_{xy} ne dépend que de τ:

$$K_{xy}(\tau) = M\{x(t_1)\, x(t_1+\tau)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 y_2 P_2(x_1, y_2, \tau)\, dx_1\, dy_2. \quad (2.88)$$

La fonction de corrélation, de même que le coefficient de corrélation (2.61) ou le moment (2.60), évalue le degré de liaison entre les valeurs d'un processus aléatoire aux instants différents.

Par définition, $K_x(t_1, t_2) = K_x(t_2, t_1)$ et pour un processus aléatoire stationnaire $K_x(\tau) = K_x(-\tau)$.

Il existe une sous-classe des processus aléatoires stationnaires appelés *ergodiques*, pour laquelle la *moyenne d'ensemble* (c'est-à-dire l'espérance mathématique) est égale, avec une probabilité unité, à la *moyenne temporelle*.

Par exemple, la moyenne temporelle de la variable $x(t)$, donnée par l'expression

$$\hat{x} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, dt,$$

est égale à la moyenne d'ensemble (2.81). Encore, la moyenne temporelle du carré de la fonction $x(t)$

$$\hat{x^2} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x^2(t)\, dt$$

est égale à la moyenne d'ensemble (2.82). Une égalité analogue a lieu également pour les moyennes du produit

$$K_x(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 P_2(x_1, x_2, \tau)\, dx_1\, dx_2 = \\ = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, x(t+\tau)\, dt. \quad (2.89)$$

En posant $\tau = 0$ on obtient

$$K_x(0) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x^2(t)\, dt = \hat{x}^2(t). \tag{2.90}$$

Cette expression s'appelle parfois *puissance* du signal $x(t)$. Si $x(t)$ est une tension appliquée à la résistance d'un ohm, $\hat{x}^2(t)$ est alors la moyenne de la puissance dégagée sur cette résistance.

La transformée de Fourier de la fonction d'autocorrélation $K_x(t)$ d'un processus aléatoire stationnaire

$$S_x(\omega) = \int_{-\infty}^{\infty} K_x(\tau)\, e^{-j\omega\tau}\, d\tau \tag{2.91}$$

s'appelle *densité spectrale* du processus aléatoire $x(t)$. La transformée de Fourier de la fonction $K_{xy}(\tau)$ (2.86), c'est-à-dire

$$S_{xy}(\omega) = \int_{-\infty}^{\infty} K_{xy}(\tau)\, e^{-j\omega\tau}\, d\tau, \tag{2.92}$$

est dite *densité spectrale mutuelle* des processus $x(t)$ et $y(t)$.

Voici l'interprétation physique de la fonction $S_x(\omega)$. Prenons sa transformée de Fourier inverse

$$K_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega)\, e^{j\omega\tau}\, d\omega. \tag{2.93}$$

Posons $\tau = 0$. Il vient

$$K_x(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega)\, d\omega. \tag{2.94}$$

Le premier membre de cette expression est la puissance du signal $x(t)$ [cf. (2.90)]. Par conséquent le second membre de (2.94) est également une puissance mais mise sous la forme d'une intégrale par rapport aux pulsations ω. La différentielle $\frac{S_x(\omega)d\omega}{2\pi}$ a le sens de la fraction de la puissance du signal correspondant à un intervalle infiniment étroit du spectre de fréquences entre ω et $\omega + d\omega$.

Le signal aléatoire dont $S(\omega) = S_0 = \text{const}$ s'appelle *bruit blanc*. L'original qui correspond à cette transformée de Fourier s'écrit

$$K_x(\tau) = S_0 \cdot \delta(\tau), \tag{2.95}$$

où $\delta(\tau)$ est ce qu'on appelle impulsion unité ou fonction de Dirac définie comme suit :

$$\left.\begin{aligned} &\delta(\tau) = 0, \quad \tau \neq 0, \\ &\int_{-\infty}^{\infty} \delta(\tau) . d\tau = 1. \end{aligned}\right\} \tag{2.96}$$

La fonction $\delta(\tau)$ est une impulsion « infiniment élevée » et « infiniment étroite » apparaissant à l'instant $\tau = 0$.

Les formules (2.95) et (2.96) montrent que pour $\tau \neq 0$ la fonction d'autocorrélation du bruit blanc est égale à zéro. On en tire qu'à des instants différents la corrélation entre les valeurs du bruit blanc n'existe pas.

Dans ce qui suit nous allons souvent remplacer un processus aléatoire continu $x(t)$ par une suite de variables aléatoires liées $x(t_1), \ldots, x(t_n)$, valeurs du processus à des instants discrets $t_1, \ldots, t_n$. Ceci est possible si le spectre de fréquences de toutes les réalisations du processus aléatoire est borné par la limite supérieure ω_s alors que la durée l'est par le temps T. En toute rigueur, le spectre de fréquences d'un processus de durée finie est infini ; on peut trouver pourtant une pulsation ω_s telle que pour $\omega > \omega_s$ l'intensité de la « queue » du spectre soit suffisamment petite et son influence soit négligeable. D'après le théorème de V. Kotelnikov [5.16] la fonction $x(t)$ peut être remplacée, sans perdre aucune information sur cette fonction, par une suite de ses valeurs discrètes $x(t_1), x(t_2), \ldots$, dites également *échantillons*, et distantes dans le temps pas plus de $\Delta t = \frac{\pi}{\omega_s}$. Le nombre total d'échantillons obtenu est égal à $\frac{T}{\Delta t} = \frac{T}{\pi}\omega_s = 2Tf_s$, où $f_s = \frac{\omega_s}{2\pi}$.

Introduisons la notation $x(t_i) = x_i$ $(i = 1, \ldots, k)$ et considérons le vecteur

$$X = (x_1, x_2, \ldots, x_k). \tag{2.97}$$

Les coordonnées du vecteur X sont des *valeurs successives* du processus $x(t)$ aux instants $t_1, t_2, \ldots, t_k$. On peut considérer le vecteur X dans un espace de dimension k aux coordonnées cartésiennes $x_1, x_2, \ldots, x_k$. Le vecteur X est aléatoire du fait que ses coefficients sont des variables aléatoires. Désignons par

$$P(X) = P(x_1, x_2, \ldots, x_k) \tag{2.98}$$

la densité de probabilité de ce vecteur ; ce n'est rien d'autre que la densité de probabilité conjointe des coordonnées $x_1, \ldots, x_k$.

L'expression $P(X)\, d\Omega(X)$, où $d\Omega(X) = dx_1, \ldots, dx_k$, donne la probabilité pour l'extrémité du vecteur X de tomber dans le volume infiniment petit $d\Omega(X)$. Comme l'extrémité du vecteur X

se trouve toujours en un point quelconque, on obtient :

$$\int_{\Omega(X)} P(X)\, d\Omega(X) = 1. \tag{2.99}$$

$\Omega(X)$ est ici l'espace de dimension k tout entier du vecteur X. C'est le cas où le processus aléatoire est remplacé par le vecteur aléatoire X.

§ 2. Méthodes variationnelles

La recherche de l'extremum d'une fonctionnelle se rapporte au domaine des problèmes dits variationnels dont le rôle dans la théorie des systèmes optimaux est évident.

Il existe diverses méthodes de résolution des problèmes de ce type. Le calcul variationnel classique a pris son essor dès 1696 lorsque Jean Bernoulli a énoncé son problème sur la ligne de plus grande pente (brachistochrone). Au XVIIIe siècle Euler et Lagrange ont établi les méthodes générales de résolution des problèmes variationnels. Plusieurs ouvrages brillants du XIXe siècle ont achevé de mettre au point le calcul variationnel classique. Le présent paragraphe passe en revue ses principes fondamentaux. Le lecteur désireux d'approfondir ses connaissances dans ce domaine peut s'adresser aux références [2.4 à 2.8].

Le XXe siècle a été marqué par l'application des méthodes dites directes de résolution des problèmes qui sont dues également à Euler. Ces derniers temps ces méthodes interviennent de plus en plus en physique et en technique. Dans ce qui suit nous en donnons un exposé sommaire (cf. également [2.7 et 2.8]).

De nouveaux problèmes, formulés vers 1950, parmi lesquels ceux de la théorie des systèmes optimaux de la commande automatique, ont nécessité l'apparition de nouvelles méthodes de résolution des problèmes variationnels ; ce sont la programmation dynamique établie par le mathématicien américain R. Bellman et ses collaborateurs et le principe du maximum énoncé et développé par l'académicien soviétique L. Pontriaguine et ses élèves.

Pour formuler le problème le plus simple du calcul variationnel considérons la fonctionnelle I associée à la fonction $y(x)$:

$$I = \int_{x_0}^{x_1} F(x, y, y')\, dx, \tag{2.100}$$

où F est une fonction donnée des arguments x, y et $y' = \frac{dy}{dx}$. Admettons d'abord que les limites d'intégration x_0 et x_1 soient des constantes données. Soit la fonction F univoque et continue avec ses dérivées partielles jusqu'au troisième ordre inclus pour tous x et y appartenant à un certain domaine R du plan (x, y). Convenons que la fonc-

tion $y = f(x)$ soit univoque et continue dans l'intervalle (x_0, x_1) et que dans cet intervalle sa dérivée première soit continue, ou, d'une façon concise, appartienne à la classe $C^{(1)}$. Appelons les courbes $f(x)$ admissibles si elles appartiennent à la classe $C^{(1)}$, reposent entièrement dans le domaine R et passent par les points donnés (x_0, y_0), (x_1, y_1), où $y_0 = f(x_0)$ et $y_1 = f(x_1)$.

Le problème se formule comme suit : rechercher celle des courbes admissibles $f(x)$ dont l'intégrale (2.100) a une valeur minimale.

La condition nécessaire que doit justifier une courbe présentant la solution du problème énoncé s'obtient facilement. Soit $f(x)$ cette courbe. Remplaçons cette fonction dans l'intégrale (2.100) par une fonction « voisine »

$$y = f(x) + \alpha\eta(x), \tag{2.101}$$

où $\eta(x)$ est une fonction arbitraire de la classe $C^{(1)}$ qui s'annule aux extrémités de l'intervalle considéré

$$\eta(x_0) = \eta(x_1) = 0, \tag{2.102}$$

et α un petit nombre quelconque. L'intégrale I devient alors fonction $I(\alpha)$ de ce nombre :

$$I(\alpha) = \int_{x_0}^{x_1} F[x, f(x) + \alpha\eta(x), f'(x) + \alpha\eta'(x)]\, dx. \tag{2.103}$$

Si α est suffisamment petit, $I(\alpha)$ peut être développée en série de puissances de α :

$$I(\alpha) = [I(\alpha)]_{\alpha=0} + \frac{\alpha}{1!}\left(\frac{dI}{d\alpha}\right)_{\alpha=0} + \frac{\alpha^2}{2!}\left(\frac{d^2I}{d\alpha^2}\right)_{\alpha=0} + \ldots \tag{2.104}$$

Les expressions $\alpha\left(\frac{dI}{d\alpha}\right)_{\alpha=0}$ et $\alpha^2\left(\frac{d^2I}{d\alpha^3}\right)_{\alpha=0}$ s'appellent variations première et seconde de l'intégrale I et sont notées δI et $\delta^2 I$ respectivement.

Si la fonction $f(x)$ minimise l'intégrale I, on a

$$\left(\frac{dI}{d\alpha}\right)_{\alpha=0} = 0, \tag{2.105}$$

de plus, cette condition se vérifie pour une fonction $\eta(x)$ quelconque de la classe $C^{(1)}$ justifiant les conditions aux limites (2.102).

Développons l'expression de δI. La dérivation de (2.103) sous le signe somme par rapport à α en posant $\alpha = 0$ donne

$$\left[\frac{dI(\alpha)}{d\alpha}\right]_{\alpha=0} = \int_{x_0}^{x_1} \left[\frac{\partial F}{\partial y}\eta(x) + \frac{\partial F}{\partial y'}\eta'(x)\right] dx. \tag{2.106}$$

Supposant que la dérivée seconde $f''(x)$ de la fonction $f(x)$ soit continue (des raisonnements plus poussés montrent que cette hypo-

thèse n'est pas impérative), le deuxième terme de (2.106) peut être intégré par parties:

$$\int_{x_0}^{x_1} \frac{\partial F}{\partial y'} \eta'(x)\, dx = \left[\eta(x) \frac{\partial F}{\partial y'}\right]_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta(x) \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) dx. \quad (2.107)$$

En vertu de (2.102), le premier terme de cette expression s'annule. Par conséquent la combinaison de (2.106) et de (2.107) conduit à la formule

$$\left[\frac{dI(\alpha)}{d\alpha}\right]_{\alpha=0} = \int_{x_0}^{x_1} \eta(x) \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right] dx. \quad (2.108)$$

Si d'après (2.105) le premier membre de (2.108) est nul, l'intégrale du deuxième membre de (2.108) est nulle également et, de plus, elle l'est pour *toute* fonction $\eta(x)$ de la classe $C^{(1)}$. On montre aisément (cf. [2.5 à 2.8]) qu'il n'en est ainsi que si l'expression entre crochets sous le signe somme est égale à zéro:

$$\frac{\partial F}{dy} - \frac{d}{\partial x}\left(\frac{\partial F}{\partial y'}\right) = 0. \quad (2.109)$$

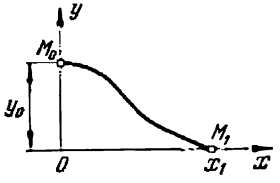


Fig. 2.4.

C'est là précisément la condition *nécessaire* qui a la forme d'une équation différentielle. En recherchant ses solutions nommées *trajectoires optimales*, on obtient les courbes qui peuvent fournir la solution du problème. L'équation (2.109) porte le nom d'Euler.

Supposons, par exemple, qu'il faut trouver la courbe $y = f(x)$ de la classe $C^{(1)}$ qui passe par les points M_0 et M_1 du plan (x, y) (fig. 2.4) de coordonnées

$$\left.\begin{aligned} x_0 = 0,\ y_0 > 0, \\ x_1 > 0,\ y_1 = 0, \end{aligned}\right\} \quad (2.110)$$

ce qui minimise l'intégrale

$$I = \int_0^{x_1} \left[y^2 + T^2 \left(\frac{dy}{dx}\right)^2\right] dx, \quad (2.111)$$

où $T^2 = \text{const}$.

Dans le cas considéré

$$F = y^2 + T^2 \left(\frac{dy}{dx}\right)^2 = y^2 + T^2 (y')^2. \quad (2.112)$$

Par conséquent

$$\frac{\partial F}{\partial y} = 2y, \quad \frac{\partial F}{\partial y'} = 2T^2 y'. \tag{2.113}$$

L'équation d'Euler (2.109) prend donc la forme

$$2y - 2T^2 \frac{d}{dx}(y') = 0, \tag{2.114}$$

ou

$$y - T^2 \frac{\partial^2 y}{\partial x^2} = 0. \tag{2.115}$$

La solution de cette équation s'écrit

$$y = C_1 e^{\frac{x}{T}} + C_2 e^{-\frac{x}{T}}, \tag{2.116}$$

où C_1 et C_2 sont des constantes. En portant dans l'expression (2.116) les conditions aux limites (2.110) on obtient la valeur des constantes

$$C_1 = \frac{y_0}{1 + e^{\frac{2x_0}{T}}}, \quad C_2 = \frac{y_0}{1 + e^{-\frac{2x_0}{T}}}. \tag{2.117}$$

Par conséquent la solution de l'équation d'Euler est donnée par la fonction

$$y = y_0 \left(\frac{e^{\frac{x}{T}}}{1 + e^{\frac{2x_0}{T}}} + \frac{e^{-\frac{x}{T}}}{1 + e^{-\frac{2x_0}{T}}} \right). \tag{2.118}$$

C'est la solution unique de l'équation (2.115) qui vérifie les conditions aux limites (2.110). Mais nous avons montré ci-dessus que s'il existe dans la classe des courbes $C^{(1)}$ une solution du problème variationnel, cette solution doit être recherchée parmi celles de l'équation d'Euler. C'est pourquoi (sous l'hypothèse qu'une solution existe dans la classe des courbes $C^{(1)}$), la courbe (2.118) intervient en donnant à l'intégrale (2.111) une valeur stationnaire. Autrement dit, l'intégrale $I(\alpha)$ peut avoir soit une valeur maximale, soit une valeur minimale, soit encore un point d'inflexion en $\alpha = 0$. Ce n'est que par raisonnements complémentaires qu'on établit réellement si (2.118) minimise l'intégrale (2.111) et donc constitue la solution du problème. Mais il ne faut pas en tirer que l'équation d'Euler, même lorsqu'elle a une solution, résout toujours le problème variationnel respectif. Il arrive que l'équation d'Euler n'a pas de solutions. Il en est ainsi, notamment, lorsque la fonction F de (2.100) ne dépend que de x et de y. L'équation d'Euler se met alors sous la forme

$$\frac{\partial F(x, y)}{\partial y} = 0. \tag{2.119}$$

Ce n'est déjà plus une équation différentielle. Par exemple, si $F(x, y) = xy$, l'équation (2.119) s'écrit $x = 0$. Recherchant la solution de l'équation (2.119) pour le cas général on peut trouver une ou plusieurs courbes $y = f(x)$. Pourtant, elles ne passent pas obligatoirement par les points imposés (x_0, y_0) et (x_1, y_1). C'est pourquoi dans le cas considéré la trajectoire optimale de la classe $C^{(1)}$ vérifiant les conditions du problème ne peut s'obtenir qu'avec des valeurs exceptionnelles des coordonnées des points frontière.

Dans d'autres cas il se peut qu'un problème variationnel de la classe $C^{(1)}$ n'ait pas de solution. Minimisons, par exemple, l'intégrale

$$I = \int_{-1}^{1} y^2 \left[1 - \left(\frac{dy}{dx}\right)\right]^2 dx \tag{2.120}$$

pour des conditions aux limites $y(-1) = 0$, $y(1) = 1$. On voit aisément que sous le signe somme la fonction $F \geqslant 0$ et, de plus, le minimum absolu de cette fonction $F = 0$ et, par conséquent, le minimum de l'intégrale I s'obtiennent sur la ligne brisée $y = 0$ avec $x \leqslant 0$, $y = x$ avec $x > 0$. Toutefois, cette ligne brisée n'appartient pas à la classe $C^{(1)}$. En l'arrondissant près d'un sommet on voit que, quelles que soient les courbes de la classe $C^{(1)}$, l'intégrale (2.120) est positive, bien que sa valeur puisse devenir aussi petite que l'on veut.

Ce qui vient d'être dit montre que la proposition: « la solution d'un problème variationnel se ramène à la résolution d'une équation d'Euler » n'est vraie que sous certaines réserves, même pour la classe restreinte des problèmes considérés.

Si l'intégrale I dépend de plusieurs fonctions d'une même variable x, il est possible par un procédé analogue à celui exposé ci-dessus de trouver la condition nécessaire pour obtenir les courbes présentant la solution du problème variationnel sous la forme d'une équation d'Euler. Soit, par exemple,

$$I = \int_{x_0}^{x_1} F(x, y, z, y', z')\, dx, \tag{2.121}$$

où $y' = \frac{dy}{dx}$, $z' = \frac{dz}{dx}$. Les valeurs frontière $y(x_0)$, $y(x_1)$, $z(x_0)$, $z(x_1)$ sont données. Considérant provisoirement que z est une fonction fixée, on obtient au lieu de (2.121) une intégrale du type (2.100) qui ne dépend que de la fonction $y(x)$. Dans ce cas la fonction réalisant l'extremum doit vérifier l'équation de la forme (2.109):

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0. \tag{2.122}$$

En rendant $y(x)$ fixe et en raisonnant d'une manière analogue par rapport à z, on peut aboutir à une autre équation

$$\frac{\partial F}{\partial z} - \frac{\partial}{dx}\left(\frac{\partial F}{dz'}\right) = 0. \tag{2.123}$$

Il se trouve finalement que les fonctions y et z doivent justifier l'ensemble des équations (2.122) et (2.123). Les fonctions cherchées $y(x)$ et $z(x)$ se trouvent parmi les solutions communes de ces équations.

Le problème variationnel peut être généralisé au cas d'une fonction comportant des dérivées d'ordres supérieurs. Convenons que la fonction $y = f(x)$ appartient à la classe $C^{(n)}$ si, avec ses dérivées jusqu'à l'ordre n inclus, elle est univoque et continue. Supposons que la fonction $y = f(x)$ de la classe $C^{(n)}$ et ses dérivées jusqu'à l'ordre $(n-1)$ aient les valeurs données lorsque $x = x_0$ et $x = x_1$, c est-à-dire

$$\left.\begin{aligned} y = y_0,\ y' = y'_0,\ \ldots,\ y^{(n-1)} = y_0^{(n-1)} \text{ avec } x = x_0, \\ y = y_1,\ y' = y'_1,\ \ldots,\ y^{(n-1)} = y_1^{(n-1)} \text{ avec } x = x_1. \end{aligned}\right\} \tag{2.124}$$

Cherchons une fonction $y = f(x)$ de ce type telle qu'elle minimise l'intégrale

$$I = \int_{x_0}^{x_1} F(x, y, y', \ldots, y^{(n)})\, dx. \tag{2.125}$$

Soit la fonction F possédant les dérivées partielles jusqu'à l'ordre $(n+2)$.

Remplaçons y par une fonction $y = \alpha\eta(x)$ voisine de la fonction cherchée, où $\eta(x)$ est une fonction de la classe $C^{(n)}$, qui devient nulle avec ses $(n-1)$ dérivées aux points frontière. Après substitution dans l'intégrale (2.125), il vient

$$\left[\frac{dI(\alpha)}{d\alpha}\right]_{\alpha=0} = \int_{x_0}^{x_1} \left[\frac{\partial F}{dy}\eta(x) + \frac{\partial F}{\partial y'}\eta'(x) + \ldots + \frac{\partial F}{\partial y^{(n)}}\eta^{(n)}(x)\right] dx. \tag{2.126}$$

Intégrons k fois par parties les termes du second membre de (2.126), il vient

$$\begin{aligned} \int_{x_0}^{x_1} \frac{\partial F}{\partial y^{(k)}}\eta^{(k)}(x)\, dx = \Big[\frac{\partial F}{\partial y^{(k)}}\eta^{(k-1)}(x) - \frac{d}{dx}\left(\frac{\partial F}{\partial y^{(k)}}\right)\eta^{(k-2)}(x) + \\ + \ldots + (-1)^{(k-1)}\frac{d^{(k-1)}}{dx^{(k-1)}}\left(\frac{\partial F}{\partial y^{(k)}}\right)\eta(x)\Big]_{x_0}^{x_1} + \\ + (-1)^k \int_{x_0}^{x_1} \frac{d^k}{dx^k}\left(\frac{\partial F}{\partial y^{(k)}}\right)\eta(x)\, dx. \end{aligned} \tag{2.127}$$

Le premier terme du second membre est nul en vertu des conditions imposées à $\eta(x)$ aux points frontière. En introduisant l'expression obtenue dans (2.126), il vient

$$\left[\frac{dI(\alpha)}{d\alpha}\right]_{\alpha=0} = \int_{x_0}^{x_1} \eta(x) \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \ldots + \right.$$

$$\left. + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial F}{\partial y^{(n)}}\right)\right] dx. \quad (2.128)$$

Etant donné que $\left(\frac{dI}{d\alpha}\right)_{\alpha=0} = 0$ pour une fonction $\eta(x)$ *quelconque* du type indiqué ci-dessus, on peut montrer (cf. [2.8]) que l'expression entre crochets sous le signe somme doit être nulle:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \ldots + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial F}{\partial y^{(n)}}\right) = 0. \quad (2.129)$$

L'équation (2.129) s'appelle équation d'Euler-Poisson. Si la fonction de la classe $C^{(n)}$ minimise l'intégrale (2.125), elle vérifie l'équation (2.129).

Cherchons (cf. [1.9]) une fonction minimisant l'intégrale

$$I_V = \int_0^\infty V\,dt = \int_0^\infty \left[x^2 + \gamma_1\left(\frac{dx}{dt}\right)^2 + \ldots + \gamma_{n-1}\left(\frac{d^{n-1}x}{dt^{n-1}}\right)^2\right] dt, \quad (2.130)$$

x étant ici un processus transitoire d'un système linéaire stable aux conditions initiales

$$(x)_{t=0} = x(0), \quad \left(\frac{dx}{dt}\right)_{t=0} = x^{(1)}(0), \ \ldots, \ \left(\frac{d^{n-1}x}{dt^{n-1}}\right)_{t=0} = x^{(n-1)}(0). \quad (2.131)$$

Le système étant stable,

$$x \to 0 \text{ et } x^{(i)} = \frac{d^i x}{dt^i} \to 0 \text{ avec } t \to \infty \ (i \quad 1, \ldots, n-1) \quad (2.132)$$

et

$$x(\infty) = x^{(1)}(\infty) = \ldots = x^{(n-1)}(\infty) = 0. \quad (2.133)$$

Les conditions (2.131) et (2.133) peuvent être considérées comme limites pour $t = 0$ et $t = \infty$.

En remplaçant dans l'équation d'Euler-Poisson y par x et x par t, on a

$$\frac{\partial F}{\partial x} - \frac{d}{dt}\left[\frac{\partial F}{\partial x^{(1)}}\right] + \ldots + (-1)^n \frac{d^n}{dt^n}\left[\frac{\partial F}{\partial x^{(n)}}\right] = 0. \quad (2.134)$$

Dans l'exemple considéré

$$F = x^2 + \gamma_1 [x^{(1)}]^2 + \ldots + \gamma_{n-1} [x^{(n-1)}]^2. \quad (2.135)$$

En portant (2.135) dans (2.134) on aboutit à l'équation différentielle

$$x - \gamma_1 x^{(2)} + \gamma_2 x^{(4)} + \ldots + (-1)^{n-1} \gamma_{n-1} x^{(2[n-1])} = 0. \quad (2.136)$$

C'est une équation linéaire d'ordre $2(n-1)$ à coefficients constants. Composons pour elle l'équation caractéristique

$$H(p) = 1 - \gamma_1 p^2 + \gamma_2 p^4 + \ldots + (-1)^{n-1} \gamma_{n-1} p^{2(n-1)} = 0. \quad (2.137)$$

Supposons que cette équation ait une racine $p_1 = \alpha_1 + j\beta_1$. Les puissances de p figurant dans l'équation (2.137) n'étant que paires, il y a également parmi ses racines $-p_1 = -\alpha_1 - j\beta_1$. On en déduit que toutes les racines de (2.137) forment des couples symétriques par rapport à l'origine des coordonnées (fig. 2.5). Donc, si parmi les racines de l'équation il n'y en a pas de purement imaginaires, la moitié, c'est-à-dire les $n-1$ racines, tombent dans le demi-plan gauche tandis que l'autre moitié, dans le demi-plan droit. Par conséquent le polynôme caractéristique peut être mis sous la forme d'un produit

$$H(p) = M(p)\, N(p). \quad (2.138)$$

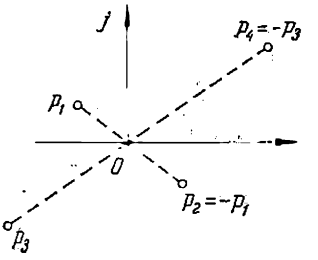


Fig. 2.5.

Notons que les racines du polynôme $M(p)$ de puissance $(n-1)$ se situent dans le demi-plan gauche, alors que celles du polynôme $N(p)$, également de puissance $(n-1)$, dans le demi-plan droit. Il est clair que seules les racines du polynôme $M(p)$ doivent être utilisées dans l'expression de la trajectoire optimale $x^*(t)$, celles du polynôme $N(p)$ se situant dans le demi-plan droit doivent être rejetées eu égard aux conditions aux limites (2.133). Autrement dit, les termes du type $C_i e^{p_i t}$ avec des p_i appartenant au demi-plan droit ne doivent pas figurer dans la solution (c'est-à-dire que les C_i respectifs doivent être nuls), sinon les conditions aux limites (2.133) seront compromises.

On peut, sur la base de certaines considérations, se donner le type de la forme quadratique V de (2.130) et chercher la trajectoire optimale correspondante. Mais on peut également s'imposer, comme c'est le cas de [1.9], l'équation de la trajectoire optimale $x^*(t)$ et rechercher l'intégrale (2.130) qui lui correspond. Soit $x^*(t)$ la solution de l'équation différentielle

$$M(D)\, x^*(t) = 0, \quad (2.139)$$

où

$$D = \frac{d}{dt} \quad (2.140)$$

est une notation symbolique de l'opération de dérivation. Donner l'équation (2.139) équivaut à donner les racines ou les coefficients de l'équation caractéristique

$$M(p) = 0. \tag{2.141}$$

Supposons ainsi que

$$M(p) = 1 + \vartheta_1 p + \vartheta_2 p^2 + \ldots + \vartheta_{n-1} p^{n-1}, \tag{2.142}$$

où $\vartheta_i > 0$. Les racines de $N(p)$ étant opposées à cel es de $M(p)$, en connaissant $M(p)$ on peut écrire le polynôme $N(p)$ (l'expression pour $N(p)$ se déduit aisément des formules de Viète):

$$N(p) = 1 - \vartheta_1 p + \vartheta_2 p^2 + \ldots + (-1)^{n-1} \vartheta_{n-1} p^{n-1}. \tag{2.143}$$

Maintenant $H(p)$ peut être donné par l'égalité (2.138). En y portant les expressions (2.142) et (2.143), on obtient les coefficients γ_i dans la formule (2.137). Par conséquent on peut trouver la forme de l'intégrale I_V (2.130) pour laquelle la trajectoire optimale (ou la famille des trajectoires optimales dans le cas des paramètres non fixés de la courbe) est la courbe donnée $x(t)$.

Considérons, par exemple, l'intégrale dont la trajectoire optimale est la solution d'une équation de deuxième ordre

$$\frac{d^2x^*}{dt^2} + 2d_0\omega_0 \frac{dx^*}{dt} + \omega_0^2 x^* = 0, \tag{2.144}$$

où $d_0 > 0$, $\omega_0 > 0$. On sait qu'avec $d_0 < 1$ la solution $x^*(t)$ de cette équation prend la forme des oscillations amorties et avec $d_0 > 1$ le processus est apériodique. La quantité ω_0 s'appelle pulsation propre des oscillations et, lorsque les valeurs de d_0 sont petites, elle est voisine de la pulsation réelle des oscillations amorties. Posons

$$\omega_0 = \frac{1}{T_0}, \quad d_0 = \sqrt{\frac{\sigma+1}{2}}. \tag{2.145}$$

L'équation (2.144) s'écrit alors

$$T_0^2 \frac{d^2x^*}{dt^2} + 2T_0 \sqrt{\frac{\sigma+1}{2}} \frac{dx^*}{dt} + x^* = 0. \tag{2.146}$$

Par conséquent les coefficients du polynôme $M(p)$ sont

$$\vartheta_1 = 2T_0 \sqrt{\frac{\sigma+1}{2}}, \quad \vartheta_2 = T_0^2. \tag{2.147}$$

Dans le cas présent l'identité (2.138) entraîne

$$(1 + \vartheta_1 p + \vartheta_2 p^2)(1 - \vartheta_1 p + \vartheta_2 p^2) = 1 - \gamma_1 p^2 + \gamma_2 p^4.$$

En ouvrant les parenthèses dans le premier membre et en égalant les coefficients de p^2 et de p^4 des premier et second membres, on tire

$$\gamma_1 = \vartheta_1^2 - 2\vartheta_2; \quad \gamma_2 = \vartheta_2^2. \tag{2.148}$$

On en déduit

$$\left.\begin{aligned} \gamma_1 &= 4T_0^2 \frac{\sigma+1}{2} - 2T_0^2 = 2\sigma T_0^2, \\ \gamma_2 &= T_0^4. \end{aligned}\right\} \tag{2.149}$$

Par conséquent l'intégrale I_V, pour laquelle la solution de l'équation (2.144) ou (2.146) est une trajectoire optimale, prend la forme

$$I_V = \int_0^\infty \{x^2 + \gamma_1 [x^{(1)}]^2 + \gamma_2 [x^{(2)}]^2\} \, dt =$$

$$= \int_0^\infty \{x^2 + 2\sigma T_0^2 [x^{(1)}]^2 + T_0^4 [x^{(2)}]^2\} \, dt. \tag{2.150}$$

Dans un grand groupe de problèmes du calcul des variations des conditions supplémentaires sont imposées à la solution. L'extremum d'une fonctionnelle, défini sous ces conditions, est dit extremum lié. Les conditions supplémentaires interviennent en général sous la forme d'égalités. Supposons, par exemple, qu'il faut chercher les courbes $y_1(x), \ldots, y_n(x)$ minimisant l'intégrale

$$I = \int_{x_0}^{x_1} F(x;\ y_1, y_2, \ldots, y_n;\ y_1', y_2', \ldots, y_n') \, dx \tag{2.151}$$

en présence des conditions supplémentaires

$$\varphi_i(x;\ y_1, y_2, \ldots, y_n) = 0 \quad (i = 1, \ldots, m;\ m < n). \tag{2.152}$$

Les équations (2.152) sont supposées indépendantes.

Pour trouver la solution on fait appel aux multiplicateurs de Lagrange. Composons l'intégrale

$$I^* = \int_{x_0}^{x_1} \left[F + \sum_{i=1}^{m} \lambda_i(x) \varphi_i\right] dx - \int_{x_0}^{x_1} F^* \, dx, \tag{2.153}$$

où

$$F^* = F + \sum_{i=1}^{m} \lambda_i(x) \varphi_i, \tag{2.154}$$

alors que $\lambda_i(x)$ sont pour le moment des fonctions inconnues (multiplicateurs de Lagrange). On recherche l'extremum absolu pour l'intégrale I^*, c'est-à-dire on résout le système d'équations d'Euler, analogue au système d'équations (2.122) et (2.123) :

$$F^*_{y_i} - \frac{d}{dx} F^*_{y_j'} = 0 \quad (j = 1, \ldots, n). \tag{2.155}$$

Si ce système de n équations est complété par un système de m équations (2.152), les $m+n$ équations suffisent généralement pour définir les $m+n$ fonctions inconnues $y_1, \ldots, y_n, \lambda_1, \ldots, \lambda_m$, alors que les conditions aux limites $y_j(x_0) = y_{j0}$ et $y_j(x_1) = y_{j1}$ $(j = 1, \ldots, n)$ qui doivent être compatibles avec les équations de liaisons (2.152) permettent de définir $2n$ constantes arbitraires.

Les conditions supplémentaires peuvent avoir la forme des équations différentielles (problème général de Lagrange) :

$$\varphi_i(x, y_1, \ldots, y_n, y_1', \ldots, y_n') = 0 \quad (i = 1, \ldots, m). \tag{2.156}$$

Dans ce cas la procédure de résolution reste la même. Les conditions supplémentaires peuvent avoir la forme des égalités intégrales (problème isopérimétrique)

$$\int_{x_0}^{x_1} F_i(x, y_1, \ldots, y_n, y_1', \ldots, y_n')\, dx = l_i \quad (i = 1, \ldots, m), \tag{2.157}$$

où l_i sont des constantes et m peut être inférieur, égal ou supérieur à n. Ce problème se ramène au précédent en introduisant de nouvelles

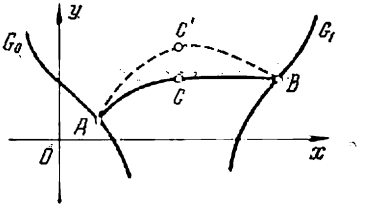


Fig. 2.6

coordonnées. La procédure d'introduction des multiplicateurs de Lagrange est ici simplifiée, car λ_i se trouvent être des constantes et il faut donc rechercher l'extremum absolu de l'intégrale

$$I^* = \int_{x_0}^{x_1} \left(F + \sum_{i=1}^{m} \lambda_i F_i\right) dx. \tag{2.158}$$

Dans des problèmes que nous avons examinés précédemment les courbes admissibles $y(x)$ étaient celles dont les extrémités se trouvaient en deux points fixés. Dans une classe plus large de problèmes les points frontière ne sont pas fixés mais doivent reposer sur des lignes ou des surfaces G_0 et G_1 données (fig. 2.6). Soit ACB la courbe recherchée. Imaginons que ses deux points frontière A et B sont fixés. Nous pouvons la comparer à d'autres courbes quelconques $AC'B$ passant par les mêmes points. Nous voyons que la

courbe ACB donne à l'intégrale I une valeur moindre que celles données par n'importe quelle courbe $AC'B$ voisine, donc, elle doit satisfaire à l'équation d'Euler. Pourtant dans la solution de l'équation d'Euler figurent des constantes arbitraires. Par exemple, les points A et B de la figure 2.6 sont définis par quatre coordonnées, tandis que les conditions d'appartenance des points A et B aux courbes frontière G_0 et G_1 ne donnent que deux égalités pour la détermination des coordonnées de ces points. L'examen détaillé des variations des points frontière que nous omettons conduit à conclure que ces points doivent satisfaire encore aux *conditions de transversalité*. Ecrivant ces conditions on peut trouver les relations qui manquent et définir les constantes des solutions des équations d'Euler.

Le XXe siècle a été marqué par une extension importante des méthodes directes du calcul des variations (cf. [2.7, 2.8]). Nous allons examiner à titre illustratif le principe de la méthode de Riesz, l'une des plus simples variantes de ces méthodes. Les fonctions considérées comme admissibles sont dans ces conditions du type suivant :

$$y_n = \sum_{i=1}^{n} \alpha_i P_i(x), \tag{2.159}$$

où α_i sont des coefficients constants et P_i certaines fonctions données. Portons l'expression (2.159) dans la formule de l'intégrale I, cette dernière se transforme alors en fonction des coefficients α_i

$$I = I(\alpha_1, \ldots, \alpha_n). \tag{2.160}$$

Maintenant on peut choisir les coefficients α_i de façon à minimiser I, en résolvant par exemple le système d'équations (non différentielles)

$$\frac{\partial I}{\partial \alpha_i} = 0 \quad (i = 1, \ldots, n). \tag{2.161}$$

Lorsque cette procédure peut être réalisée pour un n quelconque, alors pour $n \to \infty$, si la limite existe, on tombe sur la fonction $y = \lim y_n$, qui avec certaines restrictions supplémentaires est la solution exacte du problème variationnel. La fonction y_n avec n suffisamment grand est une solution approchée du problème.

Considérons maintenant les possibilités que présentent les méthodes décrites ci-dessus pour la résolution des problèmes variationnels de la théorie des systèmes optimaux. Ces problèmes comparés à ceux examinés par le calcul des variations classique permettent de relever les particularités suivantes :

1) Dans l'intégrale à minimiser, ou la fonctionnelle Q, ainsi que dans les équations de l'objet commandé et les conditions de restriction figurent, outre les coordonnées x_i de l'objet, les commandes u_j $(j = 1, \ldots, r)$.

2) Les restrictions ont généralement la forme des inégalités, par exemple $|u_j| \leqslant U_j$, et de plus le vecteur $\bar{u}$ peut se trouver non seulement *à l'intérieur* mais aussi *à la frontière* de son domaine admissible $\Omega(\bar{u})$.

3) Il arrive souvent que la solution du problème optimal est donnée par des fonctions continues par morceaux $u_j(t)$ à nombre fini de points de discontinuité de première espèce, les instants où se produisent des sauts u_j n'étant pas définis d'avance (la recherche de ces instants constitue souvent la solution du problème, cf. chapitre III).

La première des particularités mentionnées ne présente aucune difficulté. Il ne faut qu'inclure u_j en tant que fonctions examinées au même titre que x_i. On tombe alors dans le cas où l'espace de phase à n dimensions des vecteurs $\bar{x}$ est remplacé par un espace à $(n+r)$ dimensions de coordonnées $x_1, \ldots, x_n, u_1, \ldots, u_r$.

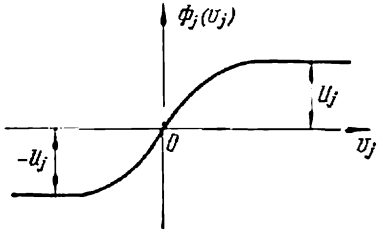


Fig. 2.7.

Les équations du système peuvent être considérées maintenant comme des conditions de restriction

$$\varphi_i = x_i' - f_i(x_1, \ldots, x_n, u_1, \ldots, u_r; t) = 0 \qquad (2.162)$$

du type (2.156). Parfois on peut exclure $\bar{u}$ en remplaçant dans l'intégrale Q et les restrictions son expression donnée par les équations (2.162), ou éliminer $\bar{x}$ s'il est possible de l'exprimer par $\bar{u}$.

La deuxième particularité donne lieu à des difficultés plus grandes. Il est vrai que dans les restrictions on peut passer des inégalités aux égalités. En effet, au lieu de u_j on peut introduire d'autres fonctions v_j associées à u_j par des égalités

$$u_j = \Phi_j(v_j) \qquad (2.163)$$

et, de plus, Φ_j peuvent être choisies telles que, quelles que soient v_j, les fonctions u_j ne sortent pas des intervalles imposés. S'il faut vérifier les conditions

$$|u_j| \leqslant U_j, \qquad (2.164)$$

on peut choisir, par exemple,

$$\Phi_j = U_j \sin v_j \qquad (2.165)$$

ou prendre Φ_j comme indiqué à la figure 2.7. La substitution (2.165) a été appliquée aux systèmes continus par Desoer [2.36], alors que la fonction du type représenté sur la figure 2.7, par Miele [2.32].

La fonction (2.165) a été appliquée aux systèmes continus à état discret par I. Tsypkine [3.30[et [3.31].

Si l'on porte maintenant $u_j = \Phi_j(v_j)$ dans l'équation (2.162), les nouvelles variables v_j peuvent être quelconques et l'on n'a plus besoin de tenir compte des restrictions (2.164) sous la forme explicite. Ce procédé permet de résoudre certains problèmes. Une condition sous la forme de l'inégalité $L \leqslant 0$ peut être remplacée généralement par une condition $M = 0$, où M est nul lorsque L est négatif et

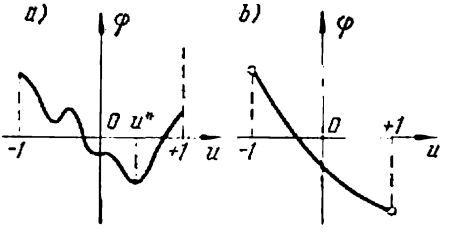


Fig. 2.8.

s'écarte du zéro quand L est positif. Pourtant, dans le cas général, l'introduction des fonctions non linéaires associées à cette transformation peut compliquer considérablement la résolution du problème.

La circonstance à première vue peu importante, que d'après (2.164) u_j peut se trouver non seulement à l'intérieur, mais aussi *à la frontière* d'un domaine fermé admissible, peut conduire parfois à de graves difficultés. Illustrons ceci sur un exemple très simple. La figure 2.8,*a* montre une fonction dérivable continue $\varphi(u)$ qui admet un minimum *à l'intérieur* de l'intervalle $|u| \leqslant 1$. Il est clair que le minimum correspond à l'un des points qui vérifient la condition

$$\frac{d\varphi}{du} = 0. \tag{2.166}$$

En résolvant l'équation (2.166) on trouve la valeur $u = u^*$ minimisant $\varphi(u)$.

Comme il faut tenir compte également des valeurs de u aux frontières de l'intervalle, il peut se trouver que le minimum repose sur une frontière (sur la figure 2.8,*b* le minimum a lieu pour $u = +1$). Il se peut alors que le point du minimum ne satisfait pas à la condition (2.166).

Enfin, la troisième particularité, comme le montre l'exposé précédent, complique sensiblement les calculs, en rendant parfois pratiquement impossible l'obtention du résultat par la méthode

classique usuelle. C'est précisément cette particularité qui est décisive du fait que dans de nombreux cas la commande optimale $\bar{u}$ admet des discontinuités de première espèce.

Il se trouve finalement que peu de problèmes de la théorie des systèmes optimaux peuvent être résolus d'une façon efficace par des méthodes classiques du calcul des variations.

Les méthodes directes ne peuvent pas non plus assurer la résolution des problèmes de commande optimale. En effet, dans de nombreux problèmes de la théorie des systèmes optimaux on ne sait pas comment choisir, sans connaître la solution, les fonctions W_i de la méthode de Riesz [cf. (2.159)] ou des méthodes analogues.

Dans ce qui suit nous allons employer de préférence des méthodes nouvelles, notamment la programmation dynamique et le principe du maximum, plus adéquates aux problèmes de la théorie des systèmes optimaux que les méthodes classiques du calcul des variations.

Il faut dire que les problèmes les plus généraux du calcul des variations, ceux de Mayer et de Boltz, sont intimement liés à la programmation dynamique et au principe du maximum (cf. [2.33, 2.34, 2.44]).

§ 3. Programmation dynamique

Dans les années cinquante le savant américain R. Bellman avec ses collaborateurs a mis au point une nouvelle méthode générale de résolution des problèmes variationnels qu'il a appelée programmation dynamique (cf. [2.9]). Par la suite, la programmation dynamique a été appliquée à une large classe de problèmes de la théorie des systèmes optimaux de commande automatique.

Revenons au problème de la commande d'un objet décrit par l'équation

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}, \bar{u}), \tag{2.167}$$

où $\bar{x}$ est le vecteur de dimension n de coordonnées $x_1, \ldots, x_n$ et $\bar{u}$ le vecteur de dimension r de coordonnées $u_1, \ldots, u_r$. Soit

$$\bar{u} \in \Phi(\bar{u}); \tag{2.168}$$

on demande de minimiser l'intégrale

$$Q = \int_0^T G[\bar{x}(t), \bar{u}(t)]\, dt, \tag{2.169}$$

où T est considéré provisoirement fixe. Nous avons noté dans le premier chapitre que le cas où G et $\bar{f}$ dépendent explicitement du temps peut être ramené aux expressions du type (2.167) et (2.169).

La programmation dynamique a à sa base le principe d'optimalité. Ce principe a été énoncé par R. Bellman pour une large classe de systèmes dont le comportement futur est défini entièrement ou statistiquement par leur état actuel. Il ne dépend donc pas de leur « préhistoire », c'est-à-dire du comportement du système dans le passé. Examinons à titre d'illustration une trajectoire optimale dans un espace de phase de dimension n (fig. 2.9) aux valeurs initiale $\bar{x}^{(0)}$ pour $t = t_0$ (généralement $t_0 = 0$) et finale $\bar{x}^{(T)}$ pour $t = T > t_0$ du vecteur $\bar{x}$. Supposons imposées les conditions initiales $\bar{x}^{(0)}$, la valeur $\bar{x}^{(T)}$ étant en général inconnue. Marquons un point intermédiaire quelconque $\bar{x}'$ de la trajectoire qui correspond à $t = t'$, où $t_0 < t' < T$, et appelons premier le segment de la trajectoire de $\bar{x}^{(0)}$ à $\bar{x}'$ et deuxième, de $\bar{x}'$ à $\bar{x}^{(T)}$. La partie de l'intégrale (2.169) correspondant au deuxième segment s'écrit $\int\limits_{t'}^{T} G[\bar{x}, \bar{u}]\, dt$. Ce deuxième segment de la trajectoire peut être considéré comme une trajectoire indépendante, qui est optimale si l'intégrale correspondante est minimale. Sur la figure 2.9 le premier et le deuxième segment sont marqués respectivement par les chiffres *1* et *2*. Dans le cas considéré le principe d'optimalité s'énonce comme suit :

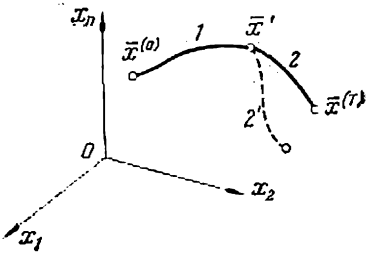


Fig. 2.9.

Le deuxième segment d'une trajectoire optimale est à son tour une trajectoire optimale.

Cela signifie que lorsque l'état initial du système est $\bar{x}'$ et l'instant initial $t = t'$, son mouvement optimal ultérieur est matérialisé par la trajectoire *2 indépendamment de la façon par laquelle le système a été amené à l'état considéré*. En effet, raisonnons par l'absurde. Le critère (2.169) étudié pour l'intervalle de temps de t' à T sera le plus petit non pas pour la trajectoire *2*, mais pour une autre trajectoire 2′ issue du point $\bar{x}'$ et représentée en pointillé sur la figure 2.9. Mais s'il en est ainsi, on pouvait construire une trajectoire « meilleure » que *1-2* pour le problème initial. Il fallait pour cela choisir la commande $\bar{u}$ telle qu'elle décrive d'abord la trajectoire *1* et puis la trajectoire 2′. Cependant, notre point de départ était que la trajectoire *1-2* est optimale. Cette contradiction prouve qu'aucune trajectoire 2′ ne peut assurer à Q une valeur inférieure à celle déterminée par la trajectoire *2*. On en déduit que la trajectoire *2* est optimale.

Le principe d'optimalité énoncé ci-dessus est une condition nécessaire, très générale, d'un processus optimal valable aussi bien pour des systèmes continus que pour des systèmes discrets.

A première vue, le principe d'optimalité est presque trivial, il semble dire peu de choses. Pourtant, comme l'a montré R. Bellman, en raisonnant avec méthode on peut en tirer les conditions nécessaires d'existence d'une trajectoire optimale qui ne sont nullement triviales. Au fond, le principe d'optimalité n'est pas du tout aussi trivial qu'il le paraît d'abord. En effet, l'affirmation qui semble être sa généralisation : « Un segment *quelconque* d'une trajectoire optimale est une trajectoire optimale », n'est pas vraie. Ainsi, le premier segment de la trajectoire $\overline{x}^{(0)}\overline{x}'\overline{x}^{(T)}$ de la figure 2.9 peut ne pas être une trajectoire optimale, c'est-à-dire qu'il peut ne pas minimiser l'intégrale (2.169) sur l'intervalle de temps de t_0 à t' si l'on impose seulement les conditions initiales $\overline{x}^{(0)}$.

Le principe d'optimalité peut être formulé d'une autre manière : *Une stratégie optimale ne dépend pas de la « préhistoire » du système, n'étant définie que par son état à l'instant considéré.*

Si l'on entend par « préhistoire » d'un système la trajectoire *1* par laquelle le point représentatif est venu en position $\overline{x}'$ (fig. 2.9), il devient clair que cette dernière formulation est équivalente à la précédente. Dans le cas considéré on entend par l'état du système à l'instant concerné l'état qui correspond au point $\overline{x}'$ à l'instant $t = t'$.

Illustrons les raisonnements de R. Bellman par un exemple simple d'un objet dont la trajectoire est définie par l'équation du premier ordre

$$\frac{dx}{dt} = f_1(x, u), \tag{2.170}$$

où x est l'unique coordonnée du système et u l'unique commande dans un certain domaine (2.168). Soit la condition initiale $x(0) = x^{(0)}$. On demande de trouver la loi de commande $u(t)$ minimisant l'intégrale

$$Q = \int_{t_0}^{T} G_1(x, u)\, dt + \varphi_1[x(T)], \tag{2.171}$$

où t_0 est généralement considéré nul et T est supposé fixe pour simplifier les raisonnements. Tout d'abord rendons le système discret, c'est-à-dire remplaçons le système continu par un système continu à état discret.

Divisons l'intervalle $(0, T)$ en N parties égales de petite longueur Δ et considérons seulement les valeurs discrètes $x = x(k)$ et $u = u(k)$ $(k = 0, 1, \ldots, N)$ aux instants $t = 0, 1\Delta, 2\Delta, \ldots, k\Delta, \ldots, (N-1)\Delta, N\Delta = T$. L'équation différentielle (2.170) de l'objet peut alors être approximativement remplacée par l'équation aux différences finies

$$\frac{x(k+1) - x(k)}{\Delta} = f_1[x(k), u(k)], \tag{2.172}$$

ou sous la forme

$$x(k+1) = x(k) + f[x(k), u(k)], \tag{2.173}$$

où

$$f[x(k), u(k)] = \Delta f_1[x(k), u(k)]. \tag{2.174}$$

La condition initiale reste la même:

$$x(0) = [x]_{t=0} = x^{(0)}. \tag{2.175}$$

L'intégrale (2.171) est remplacée approximativement par la somme

$$Q = \sum_{n=0}^{N-1} G[x(k), u(k)] + \varphi[x(N)], \tag{2.176}$$

où

$$\left.\begin{aligned} G[x(k), u(k)] &= G_1[x(k), u(k)]\,\Delta, \\ \varphi[x(N)] &= \varphi_1[x(N\Delta)] = \varphi_1[x(T)]. \end{aligned}\right\} \tag{2.177}$$

Le problème consiste maintenant à définir la suite des valeurs discrètes de la commande u, c'est-à-dire des valeurs $u(0)$, $u(1)$, ..., $u(N-1)$ qui minimisent la somme (2.176) sous les conditions (2.168), (2.173) et (2.175) imposées au système. Il s'agit ainsi de minimiser une fonction complexe de plusieurs variables. Nous verrons que la programmation dynamique permet de ramener le problème à des minimisations consécutives des fonctions d'une *seule* variable.

Pour résoudre le problème on procède par récurrence en allant de l'instant $t = T$ vers le début du processus. Considérons d'abord l'instant $t = (N-1)\Delta$. Toutes les valeurs $u(i)$ ($i = 0, 1, \ldots$ ..., $N-2$), sauf la dernière $u(N-1)$, ont été réalisées, et on a obtenu une certaine valeur $x(N-1)$ associée à l'instant $t = (N-1)\Delta$. D'après le principe d'optimalité, l'action $u(N-1)$ ne dépend pas de la « préhistoire » du système, elle n'est définie que par l'état $x(N-1)$ et le but de la commande. Considérons la dernière portion de la trajectoire de $t = (N-1)\Delta$ à $t = N\Delta$. La grandeur $u(N-1)$ n'influe que sur les termes de la somme (2.176) qui se rapportent à cette portion. Désignons la somme de ces termes par Q_{N-1}:

$$Q_{N-1} = G[x(N-1), u(N-1)] + \varphi[x(N)]. \tag{2.178}$$

(2.173) entraîne

$$x(N) = x(N-1) + f[x(N-1), u(N-1)]. \tag{2.179}$$

Par conséquent $x(N)$ dépend également de $u(N-1)$. Cherchons la valeur admissible de $u(N-1)$ vérifiant (2.168) et minimisant Q_{N-1}. Désignons par S_{N-1} la valeur minimale trouvée de Q_{N-1}. Il est clair que cette valeur est définie par l'état du système à l'instant $t = (N-1)\Delta$, c'est-à-dire par la valeur de $x(N-1)$

présente dans (2.178) et (2.179). Ainsi $S_{N-1} = S_{N-1}[x(N-1)]$. Ecrivons l'expression pour S_{N-1}:

$$S_{N-1}[x(N-1)] = \min_{u(N-1)\in\Omega(u)} Q_{N-1} =$$
$$= \min_{u(N-1)\in\Omega(u)} \{G[x(N-1),\ u(N-1)] + \varphi[x(N)]\} =$$
$$= \min_{u(N-1)\in\Omega(u)} \{G[x(N-1),\ u(N-1)] +$$
$$+ \varphi[x(N-1) + f[x(N-1),\ u(N-1)]]\}. \qquad (2.180)$$

Remarquons que pour définir S_{N-1} il ne faut minimiser que par rapport à une seule variable $u(N-1)$. Après avoir réalisé cette opération on obtient S_{N-1} sous la forme d'une fonction de $x(N-1)$; cette fonction est à retenir (par exemple, dans une mémoire quelconque si le calcul se fait avec une calculatrice numérique) avant de passer aux étapes suivantes du calcul.

Examinons maintenant l'avant-dernier intervalle du temps. En considérant ensemble le dernier et l'avant-dernier intervalles, on constate que le choix de $u(N-2)$ et de $u(N-1)$ n'influe que sur les termes de la somme (2.176) qui font partie de l'expression

$$Q_{N-2} = G[x(N-2),\ u(N-2)] +$$
$$+ \{G[x(N-1),\ u(N-1)] + \varphi[x(N)]\}. \qquad (2.181)$$

Nous pouvons considérer comme donnée la valeur $x(N-2)$ correspondant au début de l'avant-dernier intervalle et fournie par la « préhistoire » du processus. Le principe d'optimalité entraîne que seule la valeur $x(N-2)$ et le but de la commande, qui consiste à minimiser Q_{N-2}, définissent la commande optimale dans l'intervalle de temps considéré. Cherchons en utilisant $u(N-2)$ et $u(N-1)$ la valeur S_{N-2} qui est le minimum de Q_{N-2}. Mais le minimum par rapport à $u(N-1)$ du terme compris entre les accolades de l'expression (2.181) a été déjà trouvé ci-dessus pour *toute* valeur de $x(N-1)$ qui à son tour dépend de $u(N-2)$. De plus, la minimisation de Q_{N-1} a permis de trouver en même temps la valeur optimale correspondante de $u(N-1)$; notons cette valeur optimale par $u^*(N-1)$. Si l'on tient compte également du fait que le premier terme de (2.181) ne dépend pas de $u(N-2)$, on peut écrire

$$S_{N-2}[x(N-2)] = \min_{\substack{u(N-2)\in\Omega(u)\\ u(N-1)\in\Omega(u)}} Q_{N-2} =$$
$$= \min_{u(N-2)\in\Omega(u)} \{G[x(N-2),\ u(N-2)] + S_{N-1}[x(N-1)]\} =$$
$$= \min_{u(N-2)\in\Omega(u)} \{G[x(N-2),\ u(N-2)] +$$
$$+ S_{N-1}[x(N-2) + f[x(N-2),\ u(N-2)]]\},$$

puisque (2.173) entraîne

$$x(N-1) = x(N-2) + f[x(N-2),\ u(N-2)].$$

Constatons qu'ici la minimisation s'effectue également par rapport à une seule variable $u(N-2)$. De plus, on trouve $u^*(N-2)$ qui est la valeur optimale de $u(N-2)$, et la valeur S_{N-2} qui donne le minimum de la fonction Q_{N-2}. Notons que $u^*(N-2)$ comme S_{N-2} sont des fonctions de $x(N-2)$. On peut placer maintenant la fonction S_{N-2} dans la mémoire et après l'obtention de sa valeur « effacer » la fonction $S_{N-1}[x(N-1)]$ désormais inutile.

Il est important que la valeur optimale trouvée $u^*(N-2)$ minimise *toute* l'expression entre accolades de la formule pour S_{N-2} et non pas le seul terme $G[x(N-2),\ y(N-2)]$. On en tire que la stratégie qui consiste à choisir chaque $u(N-j)$ par minimisation du seul terme respectif $G[x(N-j),\ u(N-j)]$ de la somme (2.176) est loin d'être optimale. Elle n'est pas du tout « perspicace », fait déjà signalé dans ce qui précède. Une stratégie optimale tient compte du but final, c'est-à-dire de la minimisation de toute l'expression entre accolades, fonction de $u(N-j)$.

La procédure décrite de récurrence peut être poursuivie. La prise en considération du troisième intervalle à partir de la fin impose l'examen de la partie de la somme Q qui dépend de $u(N-3)$. Désignons cette partie par Q_{N-3}, on a

$$Q_{N-3} = G[x(N-3),\ u(N-3)] + \{G[x(N-2),\ u(N-2)] + \\ + G[x(N-1),\ u(N-1)] + \varphi[x(N)]\}.$$

L'expression (2.179) permet d'écrire

$$x(N-2) = x(N-3) + f[x(N-3),\ u(N-3)].$$

Ensuite, le minimum du terme entre accolades de l'expression Q_{N-3} est égal à $S_{N-2}[x(N-2)]$. Le minimum S_{N-3} de l'expression de Q_{N-3} s'écrit

$$S_{N-3}[x(N-3)] = \min_{u(N-3)\in\Omega(u)} \{G[x(N-3),\ u(N-3)] + \\ + S_{N-2}[x(N-2)]\} = \min_{u(N-3)\in\Omega(u)} \{G[x(N-3),\ u(N-3)] + \\ + S_{N-2}[x(N-3) + f[x(N-3),\ u(N-3)]]\}.$$

En passant d'une façon analogue à $S_{N-4}, \ldots, S_{N-k}$, on obtient a lformule de récurrence

$$S_{N-k}[x(N-k)] = \min_{u(N-k)\in\Omega(u)} \{G[x(N-k),\ u(N-k)] + \\ + S_{N-k+1}[x(N-k) + f[x(N-k),\ u(N-k)]]\}. \qquad (2.182)$$

Tout en effectuant la minimisation du second membre de cette formule, on définit la valeur optimale u^*, qui est fonction de

$x(N-k)$:

$$u^*(N-k) = u^*[x(N-k)] \tag{2.183}$$

et qui minimise l'expression entre accolades de (2.182).

En calculant de proche en proche d'après la formule (2.182) les S_{N-k} pour $k = 1, 2, \ldots, N$ on aboutit à la valeur optimale $u^*(0)$, c'est-à-dire à la valeur du signal de la commande, qui doit agir à l'instant initial. En fin de compte, c'est précisément cette valeur qu'il fallait connaître, car on peut toujours supposer que l'instant considéré comme courant coïncide avec l'instant initial, alors que les instants ultérieurs se rapportent au futur. Avec la valeur $u^*(0)$ on obtient aussi S_0, qui est la valeur minimale du critère Q dans le cas d'une commande optimale.

La procédure considérée peut être facilement appliquée à un objet d'ordre n quelconque décrit par l'équation (2.167) et à un nombre quelconque de commandes u_l $(l = 1, \ldots, r)$. Il suffit de remplacer les scalaires x, u et f des formules données ci-dessus par les vecteurs $\bar{x}$, $\bar{u}$ et $\bar{f}$ et d'introduire les vecteurs correspondant à l'instant $t = k\Delta$:

$$\left.\begin{aligned} \bar{x}(k) &= \{x_1(k), \ldots, x_n(k)\}, \\ \bar{u}(k) &= \{u_1(k), \ldots, u_r(k)\}. \end{aligned}\right\} \tag{2.184}$$

Ici $u_j(N-k)$ est la j-ième commande et $x_j(N-k)$ la j-ième coordonnée à l'instant $t = (N-k)\Delta$.

Remplaçons les équations différentielles (2.167) par des équations aux différences finies, et l'intégrale (2.169) par une somme. Des raisonnements analogues à ceux donnés ci-dessus montrent qu'au lieu de la formule (2.182) on a l'expression

$$S_{N-k}[\bar{x}(N-k) = \min_{\bar{u}(N-k)\in\Omega(\bar{u})} \{G[\bar{x}(N-k), \bar{u}(N-k)] + \\ + S_{N-k+1}[\bar{x}(N-k) + \bar{f}[\bar{x}(N-k), \bar{u}(N-k)]]\}. \tag{2.185}$$

La procédure du calcul ne change pas si $\bar{f}$ comprend une fonction explicite du temps.

Maintenant à chaque étape on a à minimiser une fonction de r variables $u_1(N-k), \ldots, u_r(N-k)$. Ensuite, les valeurs optimales données par le scalaire S_{N-k} et le vecteur $\bar{u}^*(N-k)$ sont les fonctions du vecteur $\bar{x}(N-k)$, c'est-à-dire les fonctions de n variables $x_1(N-k), \ldots, x_n(N-k)$.

La procédure qui vient d'être décrite peut être appliquée, sans modifications de principe, aux systèmes optimaux à processus aléatoires. Considérons, à titre d'illustration, un exemple dans lequel un objet de premier ordre subit, en plus de u, une perturbation

aléatoire z. L'équation (2.173) est alors remplacée par l'égalité

$$x(k+1) = x(k) + f[x(k), u(k), z(k)], \tag{2.186}$$

où $z(k)$ sont les valeurs discrètes de la perturbation. Maintenant $x(k)$ et le critère (2.176) deviennent des variables aléatoires. C'est pourquoi choisissons comme un nouveau critère Q, dont la valeur est à minimiser, l'espérance mathématique de l'expression (2.176); de plus, pour généraliser, introduisons z comme argument de G

$$Q = M\{\sum_{n=0}^{N-1} G[x(k), u(k), z(k)] + \varphi[x(N)]\}; \tag{2.187}$$

ici M désigne l'espérance mathématique. Supposons dans l'exemple considéré les variables $z(i)$ et $z(j)$, $i \neq j$, indépendantes et les densités de probabilité $P[z(0)]$, $P[z(1)]$, ..., $P[z(N)]$ connues. En utilisant la procédure décrite ci-dessus cherchons d'abord pour tout $x(N-1)$ fixé la fonction

$$\begin{aligned} S_{N-1}[x(N-1)] &= \min_{u(N-1)\in\Omega(u)} Q_{N-1} = \\ &= \min_{u(N-1)\in\Omega(u)} M\{G[x(N-1), u(N-1), z(N-1)] + \\ &+ \varphi[x(N-1) + f[x(N-1), u(N-1), z(N-1)]]\} = \\ &= \min_{u(N-1)\in\Omega(u)} \int_{-\infty}^{\infty} P[z(N-1)] \times \\ &\times \{G[x(N-1), u(N-1), z(N-1)] + \varphi[x(N-1) + \\ &+ f[x(N-1), u(N-1), z(N-1)]]\}\, dz(N-1). \end{aligned} \tag{2.188}$$

La minimisation fournit simultanément la valeur optimale $u^*[x(N-1)]$. Après avoir retenu $S_{N-1}[x(N-1)]$, nous recherchons la fonction

$$\begin{aligned} S_{N-2}[x(N-2)] &= \min_{u(N-2)\in\Omega(u)} M\{G[x(N-2), u(N-2), z(N-2)] + \\ &+ S_{N-1}[x(N-1)]\} = \min_{u(N-2)\in\Omega(u)} \int_{-\infty}^{\infty} P[z(N-2)] \times \\ &\times \{G[x(N-2)], u(N-2), z(N-2)] + S_{N-1}[x(N-2) + \\ &+ f[x(N-2), u(N-2), z(N-2)]]\}\, dz(N-2), \end{aligned} \tag{2.189}$$

etc. Ainsi, la méthode de calcul est en principe la même que pour les systèmes réguliers. Une procédure analogue peut être appliquée aux objets d'un ordre quelconque. On peut étudier également des problèmes plus généraux, pour lesquels $P[z(i)]$ sont inconnues d'avance et lorsqu'une certaine procédure optimale de traitement

des observations permet d'accumuler l'information sur les densités de probabilité (cf. [2.38, 5.32, 5.33]).

Sous certaines hypothèses supplémentaires, la programmation dynamique peut être appliquée à l'étude des systèmes continus. Supposons que la trajectoire de l'objet est caractérisée par les équations

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}, \bar{u}, t). \tag{2.190}$$

A l'instant initial t_0 le vecteur $\bar{x}$ est égal à $\bar{x}^{(0)}$ et le critère d'optimalité s'écrit

$$Q = \int_{t_0}^{T} G(\bar{x}, \bar{u}, t)\, dt, \tag{2.191}$$

en supposant pour simplifier $T = \text{const}$ (ce qui en général n'est pas obligatoire; cf., par exemple, chapitre IV).

Supposons que l'on ait trouvé la trajectoire optimale menant du point initial $\bar{x}^{(0)}$ au point final $\bar{x}^{(T)}$ (fig. 2.9). Notons $S(\bar{x}^{(0)}, t_0)$ la valeur minimale du critère Q correspondant à la trajectoire optimale. D'après le principe d'optimalité, la portion de la trajectoire entre le point $\bar{x}$ qui correspond à l'instant $t > t_0$ et le point final $\bar{x}^{(T)}$ (fig. 2.10) est également une trajectoire optimale, par conséquent, la partie du critère Q correspondant à cette portion et à l'intervalle de temps entre t et T a une valeur minimale possible. Désignons cette valeur par $S[\bar{x}(t), t]$.

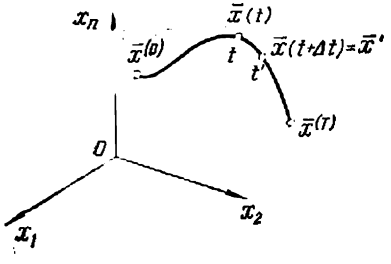


Fig. 2.10.

Soit Δt un petit intervalle de temps et $S[\bar{x}(t+\Delta t), t+\Delta t] = S[\bar{x}', t']$ la valeur minimale de la partie de l'intégrale Q qui correspond à la portion de la trajectoire optimale du point $\bar{x}(t+\Delta t) = \bar{x}'$ au point final $\bar{x}^{(T)}$ et, par conséquent, à l'intervalle de temps de $t + \Delta t = t'$ à T. La relation entre $S[\bar{x}', t']$ et $S[\bar{x}, t]$ s'exprime par la formule (2.185) à condition de substituer $S[\bar{x}, t]$ à $S_{N-k}[\bar{x}(N-k)]$, $S[\bar{x}', t']$ à $S_{N-k+1}[\bar{x}(N-k+1)]$ et enfin $G[\bar{x}(t), \bar{u}(t), t]\,\Delta t$ à $G[\bar{x}(N-k), \bar{u}(N-k)]$. Cette dernière substitution a été réalisée dans la première des équations (2.177). Etant donné que Δt est un intervalle de temps petit mais fini et que le remplacement de l'équation différentielle par une expression aux différences finies est imprécis, il faut encore ajouter à l'un des

membres de l'égalité l'expression $o_1(\Delta t)$, c'est-à-dire une grandeur d'un ordre de petitesse supérieur à celui de Δt. Cela signifie que

$$\lim_{\Delta t \to 0} \frac{o_1(\Delta t)}{(\Delta t)} = 0. \tag{2.192}$$

Ainsi, au lieu de l'égalité (2.185) on peut écrire

$$S[\bar{x}, t] = \min_{\bar{u}(t) \in \Omega(\bar{u})} \{G[\bar{x}, \bar{u}, t]\,\Delta t + S[\bar{x}', t']\} + o_1(\Delta t). \tag{2.193}$$

On pourrait obtenir l'équation (2.193) indépendamment du cas discret examiné ci-dessus. En effet, par définition

$$S[\bar{x}, t] = \min_{\bar{u}(\tau) \in \Omega(\bar{u})} \int_t^T G(\bar{x}, \bar{u}, \tau)\, d\tau \qquad (t \leqslant \tau \leqslant T). \tag{2.194}$$

Ici S est la valeur minimale de l'intégrale définie sur l'ensemble de toutes les commandes admissibles $\bar{u}(\tau)$ dans l'intervalle de t à T.

L'intégrale (2.194) peut être mise sous la forme d'une somme de deux termes associés aux intervalles $[t, t + \Delta t]$ et $[t + \Delta t, T]$. Comme Δt est petit, on peut écrire

$$S[\bar{x}, t] = \min_{\bar{u}(\tau) \in \Omega(\bar{u})} \left[G(\bar{x}, \bar{u}, t)\,\Delta t + \int_{t'=t+\Delta t}^{T} G(\bar{x}, \bar{u}, \nu)\, d\nu\right] + o_1(\Delta t), \tag{2.195}$$

où Δt est supposé petit et $o_1(\Delta t)$ d'un ordre de petitesse supérieur à celui de Δt. Remarquant que le premier terme entre crochets de (2.195) ne dépend que de la valeur de $\bar{u}(t)$ à l'instant t et que seule l'intégrale est fonction des valeurs de $u(\nu)$ dans l'intervalle de variation ν de $t' = t + \Delta t$ à T, on peut écrire

$$S[\bar{x}, t] = \min_{\bar{u}(t) \in \Omega(\bar{u})} \left[G(\bar{x}, \bar{u}, t)\,\Delta t + \min_{\bar{u}(\nu) \in \Omega(\bar{u})} \int_{t'}^{T} G(\bar{x}, \bar{u}, t)\, d\nu\right] + o_1(\Delta t) =$$
$$= \min_{\bar{u}(t) \in \Omega(\bar{u})} \{G(x, \bar{u}, t)\,\Delta t + S[\bar{x}', t']\} + o_1(\Delta t). \tag{2.196}$$

Le minimum de l'expression entre crochets se rapporte à la valeur de $\bar{u}(t)$ à l'instant t. Les formules (2.196) et (2.193) coïncident.

De même que dans la formule (2.185) il faut tenir compte du fait que $\bar{x}' = \bar{x}(t + \Delta t)$ dépend de $\bar{u}(t)$. Pour Δt petits (2.190) entraîne

$$\bar{x}' = \bar{x}(t + \Delta t) = \bar{x}(t) + \frac{d\bar{x}}{dt}\,\Delta t + o_2(\Delta t) =$$
$$= \bar{x}(t) + f[\bar{x}(t), \bar{u}(t), t]\,\Delta t + o_2(\Delta t), \tag{2.197}$$

où $o_2(\Delta t)$ est une grandeur d'ordre de petitesse supérieur à celui de Δt. La formule (2.197) est analogue à l'expression (2.173).

Supposons maintenant que S compte des dérivées partielles par rapport aux variables x_i $(i = 1, \ldots, n)$ et t, c'est-à-dire que tous les $\frac{\partial S}{\partial x_i}$ $(i = 1, \ldots, n)$ et $\frac{\partial S}{\partial t}$ existent. Pour le moment rien ne confirme cette hypothèse, cependant, la validité des résultats qui peuvent en être déduits dépend de son authenticité. Si elle ne se vérifie pas, les raisonnements qui suivent n'ont qu'un sens heuristique. Or dans certains cas cette hypothèse n'est pas vraie. C'est pourquoi pour pouvoir appliquer la programmation dynamique aux systèmes continus, il faut en général remplir certaines conditions supplémentaires (cf. [2.31, 2.33]).

Portons l'expression de $\bar{x}'$ tirée de (2.197) dans la formule (2.193) et développons $S[\bar{x}', t']$ en série de Taylor dans le voisinage du point $(\bar{x}, t)$, il vient

$$S[\bar{x}', t'] = S[\bar{x}(t+\Delta t), t+\Delta t] =$$
$$= S[\bar{x}(t) + \bar{f}[\bar{x}(t), \bar{u}(t), t]\Delta t + o_2(\Delta t); t+\Delta t] =$$
$$S[\bar{x}, t] + \sum_{i=1}^{n} \frac{\partial S[\bar{x}, t]}{\partial x_i} f_i[\bar{x}, \bar{u}, t]\Delta t + \frac{\partial S[\bar{x}, t]}{\partial t}\Delta t + o_3(\Delta t), \quad (2.198)$$

où $o_3(\Delta t)$ est d'un ordre de petitesse supérieur par rapport à Δt.

Cette formule peut être écrite sous une forme plus condensée en introduisant le gradient de la fonction $S[\bar{x}, t]$ qui est le vecteur de coordonnées $\frac{\partial S}{\partial x_i}$ $(i = 1, \ldots, n)$, soit

$$\operatorname{grad} S = \left(\frac{\partial S}{\partial x_1}, \ldots, \frac{\partial S}{\partial x_n}\right). \quad (2.199)$$

La formule (2.198) se met alors sous la forme

$$S[\bar{x}', t'] = S[\bar{x}(t+\Delta t), t+\Delta t] =$$
$$= S[\bar{x}, t] + \langle \operatorname{grad} S[\bar{x}, t] \cdot \bar{f}[\bar{x}(t), \bar{u}(t), t]\rangle \Delta t +$$
$$+ \frac{\partial S[\bar{x}, t]}{\partial t}\Delta t + o_3(\Delta t). \quad (2.200)$$

Les signes $\langle\ \rangle$ désignent ici le produit scalaire des vecteurs grad S et $\bar{f}$. Portons (2.200) dans (2.193) et mettons en facteur les grandeurs $S[\bar{x}, t]$ et $\frac{\partial S}{\partial t}$ du second membre, ne dépendant pas de $\bar{u}(t)$. Ensuite, $S[\bar{x}, t]$ des premier et second membres s'annulent

réciproquement et en divisant la formule par Δt on obtient :

$$-\frac{\partial S[\bar{x}, t]}{\partial t} = \min_{\bar{u}(t)\in\Omega(\bar{u})} \{G[\bar{x}(t), \bar{u}(t), t] + \\ + \langle \operatorname{grad} S[\bar{x}, t]\cdot \bar{f}[\bar{x}(t), \bar{u}(t), t]\rangle\} + \frac{o_4(\Delta t)}{\Delta t}, \quad (2.201)$$

où $o_4(\Delta t)$ est d'un ordre de petitesse supérieur par rapport à Δt. Maintenant faisons tendre Δt vers zéro. Puisque $o_4(\Delta t)$ vérifie une condition du type (2.192), le dernier terme du second membre de (2.201) disparaît pour $\Delta t \to 0$. Il vient donc à la limite :

$$-\frac{\partial S[\bar{x}, t]}{\partial t} = \min_{\bar{u}(t)\in\Omega(\bar{u})} \{G[\bar{x}(t), \bar{u}(t), t] + \\ + \langle \operatorname{grad} S[\bar{x}, t]\cdot \bar{f}[\bar{x}(t), \bar{u}(t), t]\rangle\}. \quad (2.202)$$

Cette expression est dite équation de Bellman. C'est une équation différentielle aux dérivées partielles, qui a ceci de particulier qu'à la suite de la minimisation, la variable $\bar{u}$ disparaît du second membre pour tout t. Considérons à titre d'illustration un exemple très simple tiré de [2.17]. Soit $r = 1$ et $n = 2$; de plus $G = G(x_1, x_2)$ et l'unique commande est notée u. Les équations de l'objet s'écrivent :

$$\frac{dx_1}{dt} = f_1 = ux_1 + x_2, \quad \frac{dx_2}{dt} = f_2 = u^2. \quad (2.203)$$

L'équation (2.202) se met alors sous la forme (pour simplifier l'écriture employons S au lieu de $S[\bar{x}, t]$) :

$$-\frac{\partial S}{\partial t} = \min_u \left\{G(x_1, x_2) + \frac{\partial S}{\partial x_1}(ux_1 + x_2) + \frac{\partial S}{\partial x_2}u^2\right\}. \quad (2.204)$$

En supposant que $\frac{\partial S}{\partial x_2} > 0$ minimisons l'expression entre accolades par rapport à u en égalant à zéro sa dérivée par rapport à u. La valeur optimale u^* minimisant l'expression entre crochets est

$$u^* = -\frac{1}{2}x_1\frac{\partial S}{\partial x_1}\cdot\frac{1}{\frac{\partial S}{\partial x_2}}. \quad (2.205)$$

En portant cette expression dans l'égalité (2.204) on a une équation aux dérivées partielles sous sa forme habituelle

$$-\frac{\partial S}{\partial t} = G(x_1, x_2) + \frac{\partial S}{\partial x_1}x_2 - x_1^2\frac{\left(\frac{\partial S}{\partial x_1}\right)^2}{4\frac{\partial S}{\partial x_2}} \quad (2.206)$$

L'équation aux dérivées partielles (2.206) peut être résolue car on connaît les conditions aux limites. En effet, $S[\bar{x}, T]$ est une fonction connue. Pour le critère (2.171), par exemple, elle est

égale à la fonction connue $\varphi_1 [x (T)]$, étant donné que pour $t_0 = T$ l'intégrale de (2.171) est égale à zéro. Pour le critère (2.191), la fonction $S [\bar{x}, T]$ est nulle. Connaissant la fonction $S [\bar{x}, T]$ on peut intégrer l'équation (2.206). L'une des méthodes d'intégration approchée usuelles consiste à trouver les valeurs discrètes de la commande et à résoudre les relations de récurrence obtenues du type (2.185). Dans certains cas on peut trouver une solution approchée par une autre méthode (cf., par exemple, [2.35]) ou même obtenir une solution exacte sous une forme fermée. Les calculs fournissent accessoirement la valeur u^* qui constitue la commande optimale.

§ 4. Principe du maximum

En 1956 l'académicien L. Pontriaguine a publié avec ses élèves V. Boltianski et R. Gamkrélidzé l'article [2.11] dans lequel il a énoncé, sous forme d'hypothèse, le principe rendant possible la résolution du problème général, qui consiste à rechercher le processus transitoire optimal du point de vue du temps de transition dans des systèmes continus. La découverte de ce principe couronne les travaux de Pontriaguine et de ses collaborateurs consacrés aux problèmes de commande optimale; certains problèmes ont été formulés par l'auteur du présent ouvrage dans les rapports sur la théorie des systèmes optimaux présentés en 1954 au séminaire dirigé par Pontriaguine. Dans les ouvrages qui depuis 1956 ont suivi la première publication [2.12-2.14, 2.16, 2.18, 2.21], Pontriaguine, Boltianski et Gamkrélidzé ont démontré que le principe du maximum est un critère nécessaire et suffisant d'un processus optimal dans les systèmes linéaires et un critère nécessaire dans les systèmes non linéaires. Par ailleurs, le principe du maximum a été généralisé au cas de la minimisation d'une intégrale et de la restriction des coordonnées de l'objet. Par la suite, L. Rozonoère a étudié d'autres méthodes [2.15, 2.17, 2.19]. Dans ses travaux il a mis en évidence le rapport existant entre le principe du maximum et la programmation dynamique et a justifié la validité du principe du maximum pour les systèmes linéaires continus à état discret (cf. aussi [1.22, 2.20]). Dans [3.43, 3.44] A. Boutkovski a généralisé le principe du maximum à quelques classes des équations intégrales décrivant les systèmes à paramètres répartis.

La démonstration du principe du maximum donnée dans les travaux de Pontriaguine et de ses collaborateurs n'a aucun rapport immédiat avec le principe d'optimalité de Bellman et la programmation dynamique. Toutefois, il nous paraît plus commode de déduire d'abord le principe du maximum de l'équation de Bellman [2.17, 2.20] pour montrer le lien qui existe entre ces deux principes. Dans ce qui suit la déduction du principe du maximum est donnée indépendamment de la programmation dynamique.

Ecrivons l'équation (2.202) sous une forme plus compacte. A cette fin introduisons une coordonnée supplémentaire x_{n+1} en posant $(x_{n+1})_{t=0} = 0$, dont l'équation est

$$\frac{dx_{n+1}}{dt} = f_{n+1} = 1. \tag{2.207}$$

Si à l'instant initial $t = 0$, alors $x_{n+1} = t$. Au lieu de t on peut écrire x_{n+1} et au lieu de $\frac{\partial S}{\partial t}$, $\frac{\partial S}{\partial x_{n+1}}$. Introduisons de plus la coordonnée x_0 [cf. (1.56)] d'équation

$$\frac{dx_0}{dt} = f_0 = G[\bar{x}, \bar{u}, t] = G[\bar{x}, \bar{u}, x_{n+1}], \tag{2.208}$$

avec $(x_0)_{t=0} = 0$. Le problème de minimisation de l'intégrale Q se ramène à la minimisation de la grandeur $(x_0)_{t=T} = x_0^{(T)}$.

Introduisons maintenant les vecteurs généralisés dans un espace à $(n + 2)$ dimensions :

$$\left.\begin{aligned} \tilde{x} &= (x_0, x_1, \ldots, x_n, x_{n+1}), \\ \tilde{f} &= (f_0, f_1, \ldots, f_n, f_{n+1}) \end{aligned}\right\} \tag{2.209}$$

et

$$\tilde{\psi} = \left(-1, -\frac{\partial S}{\partial x_1}, \ldots, -\frac{\partial S}{\partial x_n}, -\frac{\partial S}{\partial x_{n+1}}\right). \tag{2.210}$$

Transposons $\left(-\frac{\partial S}{\partial t}\right)$ de l'expression (2.202) dans le second membre tout en retenant que le minimum de l'expression du second membre signifie le maximum, pris avec le signe moins, de l'expression de signe opposé. En effet, pour un μ quelconque on a la relation

$$\max(-\mu) = -\min \mu. \tag{2.211}$$

Si l'on tient compte de (2.207), (2.208) et (2.211), l'expression (2.202) peut être mise sous la forme

$$0 = \max_{\bar{u}(t)\in\Omega(\bar{u})} \Big\{ G[\bar{x}, \bar{u}, x_{n+1}]\cdot(-1) - \\ - \langle \operatorname{grad} S[\bar{x}, x_{n+1}]\cdot \bar{f}[\bar{x}, \bar{u}, x_{n+1}]\rangle - \frac{\partial S}{\partial x_{n+1}}(+1)\Big\}. \tag{2.212}$$

En comparant cette relation avec les expressions (2.209) et (2.210) pour les vecteurs $\tilde{f}$ et $\tilde{\psi}$ on voit que la condition (2.212) peut s'écrire sous une forme très compacte :

$$0 = \max_{\bar{u}(t)\in\Omega(\bar{u})} \{\langle\tilde{\psi}\tilde{f}\rangle\}. \tag{2.213}$$

Introduisons maintenant ce qu'on appelle un hamiltonien ; c'est le scalaire

$$\tilde{H} = \langle\tilde{\psi}\tilde{f}\rangle = \sum_{i=0}^{n+1} \tilde{\psi}_i \tilde{f}_i, \tag{2.214}$$

où $\tilde{\psi}_i$ et $\tilde{f}_i$ sont les i-ièmes coordonnées des vecteurs $\tilde{\psi}$ et $\tilde{f}$ respectivement. L'égalité (2.213) s'écrit alors

$$0 = \max_{\bar{u}(t) \in \Omega(\bar{u})} \tilde{H}. \tag{2.215}$$

Cette expression traduit précisément le principe du maximum de Pontriaguine.

L'expression (2.215) entraîne deux conclusions :

a) Si le processus est optimal, et pour déduire (2.202) c'est précisément notre hypothèse de départ, à tout instant t, $\bar{u}^*(t)$ est une commande optimale, commande qui rend maximale la grandeur $\tilde{H}$ et, de plus,

$$\tilde{H}_{\max} = \max_{\bar{u}(t) \in \Omega(\bar{u})} \tilde{H} = \max_{\bar{u}(t) \in \Omega(\bar{u})} \langle \tilde{\psi} \tilde{f} \rangle =$$

$$= \max_{\bar{u}(t) \in \Omega(\bar{u})} \sum_{i=0}^{n+1} \tilde{\psi}_i \tilde{f}_i. \tag{2.216}$$

Dans cette formule la grandeur $\tilde{H}$ dépend de $\bar{u}$ car le vecteur $\tilde{f}$ en dépend. Dans un point considéré de l'espace $\tilde{x}$ à $(n + 2)$ dimensions la grandeur $\tilde{H}$ est bien définie comme fonction de $\bar{u}$ si l'on connaît le vecteur $\tilde{\psi}$; or ce dernier est bien défini si l'on connaît la fonction $S\ [\bar{x}, x_{n+1}]$ et, par conséquent, ses dérivées partielles $\frac{\partial S}{\partial x_i}$ $(i = 1, \ldots, n + 1)$.

Ainsi le choix de la commande optimale est en principe très simple ; il faut à chaque instant donné choisir $\bar{u}$ de sorte qu'on puisse assurer à l'hamiltonien $\tilde{H}$ une valeur maximale possible (compte tenu des restrictions imposées à $\bar{u}$).

b) La valeur maximale de la grandeur $\tilde{H}$ est la même en n'importe quel point de la trajectoire optimale : elle est nulle.

Pour donner l'interprétation géométrique introduisons la fonction auxiliaire $\tilde{S}$, qui est fonction du point $\tilde{x}$ dans un espace à $(n + 2)$ dimensions, définie par la formule

$$\tilde{S} = x_0 + S\ (x_1, \ldots, x_n, x_{n+1}). \tag{2.217}$$

Considérons dans l'espace $\tilde{x}$ à $(n + 2)$ dimensions la trajectoire du point représentatif qui se déplace de la position initiale $\tilde{x}^{(0)}$ vers la position finale $\tilde{x}^{(T)}$.

Etant donné que $\tilde{S}$ est une fonction continue d'un point de l'espace $\tilde{x}$, on peut trouver dans cet espace une surface qui constitue le lieu géométrique des points $\tilde{S}$ = const. Appelons les surfaces de ce type *isosurfaces*.

En comparant (2.217) et (2.210) on voit que les coordonnées du vecteur $\widetilde{\psi}$ sont assujetties à $\widetilde{S}$ par les égalités

$$\widetilde{\psi}_i = -\frac{\partial \widetilde{S}}{\partial \widetilde{x}_i}. \tag{2.218}$$

Par conséquent, le vecteur $\widetilde{\psi}$ est le gradient pris avec le signe moins du scalaire $\widetilde{S}$ dans l'espace $\widetilde{x}$ à $(n+2)$ dimensions :

$$\widetilde{\psi} = -\operatorname{grad} \widetilde{S}. \tag{2.219}$$

On sait de l'analyse vectorielle que le gradient est orthogonal à la surface $\widetilde{S} = \text{const}$. Considérons au point $\widetilde{x}$ le vecteur

$$\widetilde{f} = \frac{d\widetilde{x}}{dt}. \tag{2.220}$$

La condition $\widetilde{H} = \max$ coïncide avec celle du maximum du produit scalaire des vecteurs $\widetilde{\psi}$ et $\widetilde{f}$ ou, le vecteur $\widetilde{\psi}$ au point $\widetilde{x}$ concerné étant donné et indépendant de $\bar{u}$, avec la condition du *maximum de la projection du vecteur $\widetilde{f}$ sur la direction $\widetilde{\psi}$*. Ainsi le principe du maximum interprété géométriquement consiste à choisir une commande $\bar{u}$ de telle façon qu'elle maximise la projection du vecteur de la vitesse du point représentatif $\frac{d\widetilde{x}}{dt}$ sur la direction de la normale à l'isosurface en point $\widetilde{x}$ considéré. Notons que la projection est négative et que d'après (2.215) sa valeur maximale est égale à zéro.

Considérons le cas particulier lorsque le temps t ne figure pas explicitement dans l'équation de la trajectoire ni dans la fonction G et lorsque le temps T du processus transitoire doit être minimal. Dans l'équation (2.208) il faut alors poser $G = 1$; de plus, $\frac{\partial S}{\partial t} = 0$. Alors (2.212) entraîne

$$\max_{\bar{u} \in \Omega(\bar{u})} \langle -\operatorname{grad} S\,[\bar{x}, t] \cdot \bar{f}\,[\bar{x}, \bar{u}] \rangle = 1. \tag{2.221}$$

Nous insistons sur le fait que $\bar{x}$ et $\bar{f}$ sont ici des vecteurs d'un espace à n dimensions.

Posons

$$H = -\langle \operatorname{grad} S \cdot \bar{f} \rangle = \langle \bar{\psi}\, \bar{f} \rangle, \tag{2.222}$$

où le vecteur $\bar{\psi}$ à n dimensions est défini par l'expression

$$\bar{\psi} = -\operatorname{grad} S. \tag{2.223}$$

La condition (2.221) s'écrit alors

$$\max_{\bar{u} \in \Omega(\bar{u})} H = 1. \tag{2.224}$$

C'est la forme que prend le principe du maximum dans le cas particulier considéré. Maintenant on peut étudier la trajectoire du

point représentatif dans un espace $\bar{x}$ à n dimensions (fig. 2.11). La commande optimale $\bar{u}$ doit être choisie de façon à maximiser le scalaire H à chaque instant donné, la valeur maximale de H en un point quelconque de la trajectoire étant égale à 1. Comme dans le cas envisagé

$$S[\bar{x}, t] = \int_t^T 1 \cdot dt = T - t, \tag{2.225}$$

où t est l'instant courant, la grandeur S traduisant le temps nécessaire pour atteindre le point final diminue à mesure que t croît. Par conséquent, le vecteur $\bar{\psi}$, qui coïncide avec le sens de la décroissance la plus rapide de S, est dirigé vers l'« intérieur » de l'isosurface S = const (fig. 2.11) entourant le point final $\bar{x}^{(T)}$. Dans ce cas particulier les isosurfaces S = const deviennent des surfaces des temps égaux $\tau_i = T - t_i$ nécessaires pour rejoindre le point final $\bar{x}^{(T)}$. Dans [3.6] ces surfaces sont appelées surfaces des isochrones. Le principe du maximum appliqué à ce cas impose un choix de $\bar{u}$ tel que la projection de la vitesse $\frac{d\bar{x}}{dt}$ du point représentatif de l'espace de phase sur la direction $\bar{\psi}$ de la normale à l'isosurface soit maximale. Des considérations physiques intuitives mènent à la même conclusion. En effet, le déplacement *le long* d'une isosurface, par exemple τ_2 = const (fig. 2.11), ne donne aucun gain puisque le point représentatif reste toujours en position d'où au meilleur des cas il peut atteindre le point final $\bar{x}^{(T)}$ au bout des mêmes τ_2 secondes. Or, plus rapide est le cheminement du point représentatif le long de la normale à isosurface, plus vite il se déplace vers l'isosurface « suivante » $\tau_2 - \Delta t$ = const et, par conséquent, le temps nécessaire pour atteindre le point final $\bar{x}^{(T)}$ devient de plus en plus petit.

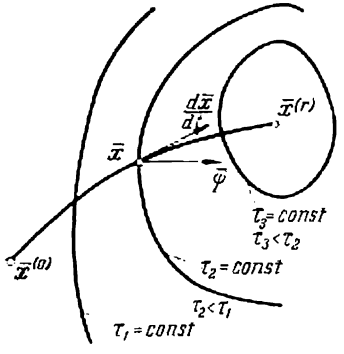


Fig. 2.11.

Nous avons dit plus haut que l'application de la méthode de programmation dynamique impose la recherche des fonctions $S[\bar{x}, t]$ ou $\hat{S}$, liée à des opérations délicates comme, par exemple, la résolution des équations aux dérivées partielles. Cependant l'emploi du principe du maximum exige qu'on connaisse seul le vecteur $\hat{\psi}$ assujetti à une trajectoire optimale et ce vecteur peut s'obtenir sans cons-

truire des surfaces $\widetilde{S} = \text{const}$. Pour trouver $\widetilde{\psi}$ il faut résoudre ce qu'on appelle des équations adjointes.

Dans les ouvrages [2.12] et [2.21] on trouve diverses méthodes de construction des équations de ce type. Nous employons une méthode analogue à celle de [2.21].

Supposons que la fonction $\widetilde{S}(\widetilde{x})$ ait des dérivées secondes partielles par rapport à tous les $\widetilde{x}_i$ $(i = 0, 1, \ldots, n+1)$, c'est-à-dire qu'elle soit deux fois dérivable. Considérons les variations du vecteur $\widetilde{\psi}$ lors du déplacement du point représentatif le long de la trajectoire optimale. Etant donné que $\widetilde{\psi} = \widetilde{\psi}[\widetilde{x}(t)]$, le vecteur $\widetilde{\psi}$ est donc une fonction du temps. On peut alors trouver les dérivées

$$\frac{d\widetilde{\psi}_i}{dt} = -\frac{d}{dt}\left(\frac{\partial \widetilde{S}}{\partial \widetilde{x}_i}\right) = -\sum_{j=0}^{n+1} \frac{\partial}{\partial \widetilde{x}_j}\left(\frac{\partial \widetilde{S}}{\partial \widetilde{x}_i}\right)\frac{d\widetilde{x}_j}{dt} =$$

$$= -\sum_{j=0}^{n+1} \frac{\partial^2 \widetilde{S}}{\partial \widetilde{x}_i \partial \widetilde{x}_j}\widetilde{f}_j \quad (i = 1, \ldots, n+1). \qquad (2.226)$$

Ici $\widetilde{f}_j$ remplace $\frac{d\widetilde{x}_j}{dt}$.

Comme on le voit de (2.210), la coordonnée $\widetilde{\psi}_0$ est toujours égale à (-1). On a donc

$$\frac{d\widetilde{\psi}_0}{dt} = 0. \qquad (2.227)$$

Si dans l'expression de $\widetilde{H}$ (2.214) on substitue à $\overline{u}(t)$ quelconque la commande optimale $u^*(t)$ le long de la trajectoire optimale considérée, on a en vertu de (2.213)

$$\widetilde{H} = \langle \widetilde{\psi}\widetilde{f}\rangle = \sum_{j=0}^{n+1} \widetilde{\psi}_j \widetilde{f}_j = -\sum_{j=0}^{n+1} \frac{\partial \widetilde{S}}{\partial \widetilde{x}_j}\widetilde{f}_j = \widetilde{H}_{\max} = 0. \qquad (2.228)$$

Fixons maintenant l'instant t, $\overline{u}^*(t)$ est alors également une grandeur fixée. Pour les points de l'espace $\widetilde{x}$ différents du point de la trajectoire concernée, la commande $\overline{u}^*(t)$ donnée n'est pas optimale ; par conséquent pour ces points la grandeur $\widetilde{H}$ n'atteint pas son maximum. Ce raisonnement entraîne qu'avec t et $u(t) = u^*(t)$ fixés, la grandeur $\widetilde{H} = \langle\widetilde{\psi}\widetilde{f}\rangle$ atteint son maximum (égal à zéro) précisément en un point de la trajectoire optimale et, par conséquent, en ce point les dérivées de $\widetilde{H}$ par rapport à $\widetilde{x}_i$ deviennent nulles. En dérivant l'expression (2.228) (tout en dérivant également $\widetilde{\psi}$ par rapport à $\widetilde{x}_i$ pour rendre compte de l'accroissement de $\widetilde{\psi}$ aux

points déplacés) on aboutit aux égalités

$$\frac{\partial}{\partial \tilde{x}_i}\left(-\sum_{j=0}^{n+1}\frac{\partial \tilde{S}}{d\tilde{x}_j}\tilde{f}_j\right)=-\sum_{j=0}^{n+1}\frac{\partial^2 \tilde{S}}{\partial \tilde{x}_j\,\partial \tilde{x}_i}\tilde{f}_j-$$

$$-\sum_{j=0}^{n+1}\frac{\partial \tilde{S}}{\partial \tilde{x}_j}\cdot\frac{\partial \tilde{f}_j}{\partial \tilde{x}_i}=0 \qquad (i=1,\ldots,n+1), \qquad (2.229)$$

d'où l'on déduit

$$-\sum_{j=0}^{n+1}\frac{\partial^2 \tilde{S}}{\partial \tilde{x}_i\,\partial \tilde{x}_j}\tilde{f}_j=\sum_{j=0}^{n+1}\frac{\partial \tilde{S}}{\partial \tilde{x}_j}\cdot\frac{\partial \tilde{f}_j}{\partial \tilde{x}_i} \qquad (i=1,\ldots,n+1). \qquad (2.230)$$

Remarquant que le premier membre de (2.230) est identique au second membre de (2.226) remplaçons le second membre de (2.226) par le second membre de (2.230). On a

$$\frac{d\tilde{\psi}_i}{dt}=\sum_{j=0}^{n+1}\frac{\partial \tilde{S}}{\partial \tilde{x}_j}\cdot\frac{\partial \tilde{f}_j}{\partial \tilde{x}_i}=-\sum_{j=0}^{n+1}\tilde{\psi}_j\frac{\partial \tilde{f}_j}{\partial \tilde{x}_i} \qquad (i=1,\ldots,n+1). \qquad (2.231)$$

Nous avons là un ensemble d'équations adjointes qui, avec (2.227), définissent la variation du vecteur $\tilde{\psi}$ sur la trajectoire optimale. Notons que les équations (2.231) sont *linéaires* par rapport aux coordonnées $\tilde{\psi}_j$ du vecteur $\tilde{\psi}$.

Dans l'expression de $\tilde{H}$ pour le point $\tilde{x}$ considéré de la trajectoire optimale $\tilde{\psi}$ est donné et il n'y a que le vecteur $\tilde{f}$ qui dépend de $\tilde{x}$ sous une forme explicite. La dérivée partielle de $\tilde{H}$ par rapport à $\tilde{x}_i$ s'écrit alors comme suit :

$$\frac{\partial \tilde{H}}{\partial \tilde{x}_i}=\sum_{j=0}^{n+1}\tilde{\psi}_j\frac{\partial \tilde{f}_j}{\partial \tilde{x}_i}. \qquad (2.232)$$

En comparant cette expression avec les équations (2.231) on remarque qu'on peut écrire ces dernières sous une forme compacte

$$\frac{d\tilde{\psi}_i}{dt}=-\frac{\partial \tilde{H}}{\partial \tilde{x}_i} \qquad (i=1,\ldots,n+1). \qquad (2.233)$$

Notons que la formule (2.228) permet de tirer

$$\frac{\partial \tilde{H}}{\partial \tilde{\psi}_i}=\tilde{f}_i, \qquad (2.234)$$

puisque $\tilde{f}$ ne dépend pas de $\tilde{\psi}_i$. Par conséquent, les équations de la trajectoire de l'objet peuvent s'écrire :

$$\frac{d\tilde{x}_i}{dt}=\frac{\partial \tilde{H}}{\partial \tilde{\psi}_i} \qquad (i=0,\ 1,\ \ldots,\ n+1). \qquad (2.235)$$

Les systèmes d'équations du type (2.233) et (2.235) sont dits équations adjointes canoniques.

Ainsi, la valeur du vecteur $\tilde{\psi}$ est donnée par les équations différentielles ordinaires (2.231) et (2.233). Mais si le vecteur $\tilde{\psi}$ est connu, nul besoin de calculer la fonction $\tilde{S}$. En appliquant le principe du maximum on résout conjointement deux systèmes d'équations, principal et adjoint.

Considérons un cas particulier où le temps n'est pas présent explicitement dans les équations de l'objet. Il s'agit d'assurer la durée minimale du processus transitoire. Dans ce cas $G = 1$ et on peut utiliser la grandeur H de (2.222) au lieu de $\tilde{H}$, ainsi que l'espace de phase $\tilde{x}$ à n dimensions au lieu de l'espace $\tilde{x}$ à $(n + 2)$ dimensions. Les équations de la trajectoire de l'objet s'écrivent alors

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial \psi_i} \quad (i = 1, \ldots, n), \tag{2.236}$$

et les équations adjointes (2.231) [cf. également (2.233)] sont de la forme

$$\frac{d\psi_i}{dt} = -\frac{\partial H}{\partial x_i} \quad (i = 1, \ldots, n). \tag{2.237}$$

Cherchons maintenant à démontrer, en nous basant sur une autre méthode, que le principe du maximum est un critère d'optimalité nécessaire. Nous allons suivre la méthode proposée par L. Pontriaguine et ses collaborateurs, mais en simplifiant légèrement l'exposé pour rendre plus aisée l'assimilation du texte. Cette méthode nous semble utile sous deux rapports. D'abord elle permet de comprendre le caractère dynamique du principe du maximum, vu sous une optique différente de celle exposée dans ce qui précède. Ensuite, indépendante de la programmation dynamique, cette méthode permet d'abandonner les hypothèses sur la dérivabilité de la fonction $S[\bar{x}, t]$ qui ne sont pas justifiées par la nature du problème.

Pour simplifier notre exposé posons que le temps T est fixé et que l'extrémité de la trajectoire est libre. Soit l'équation de la trajectoire de l'objet

$$\frac{d\tilde{x}}{dt} = \tilde{f}(\tilde{x}, \bar{u}) \tag{2.238}$$

et la grandeur à minimiser

$$Q = \tilde{x}_0^{(T)}. \tag{2.239}$$

Considérons les fonctions $\tilde{f}_i$ bornées et continues par rapport à tous leurs arguments et dérivables par rapport à $\tilde{x}_j$ $(j = 0, 1, \ldots, n + 1)$. Soit encore la commande $\bar{u}(t)$, qui appartient à la

classe des fonctions continues par morceaux vérifiant la condition

$$\bar{u}(t) \in \Omega(\bar{u}). \tag{2.240}$$

Admettons que la valeur initiale du vecteur $\tilde{x}$ soit donnée

$$(\tilde{x})_{t=0} = \tilde{x}^{(0)}. \tag{2.241}$$

Sous ces conditions il faut définir la trajectoire optimale $\tilde{x}^*(t)$ $(0 \leqslant t \leqslant T)$ et la commande optimale $\bar{u}^*(t)$ minimisant le critère Q.

Pour plus de simplicité nous examinons dans ce qui suit le cas $r = 1$ à commande unique; il est clair que les raisonnements et résultats obtenus sont valides également dans le cas $r > 1$.

Supposons que nous avons trouvé $\tilde{x}^*(t)$ et $u^*(t)$.

Considérons la courbe $u^*(t)$ (fig. 2.12). Nous avons déjà remarqué que cette courbe peut contenir un nombre fini de points aux discontinuités de première espèce. Proposons-nous de choisir la commande optimale sur un intervalle de temps infiniment petit

$$\tau - \varepsilon < t < \tau, \tag{2.242}$$

où ε est une grandeur infiniment petite, $0 < \tau < T$. Faisons varier la commande sur un intervalle infiniment petit (2.242), en la modifiant à partir de la grandeur optimale u^* jusqu'à une autre grandeur $u \neq u^*$. De plus, sur tous les autres intervalles $(0, \tau - \varepsilon)$ et (τ, T) la commande reste inchangée et égale à $u^*(t)$. La variation de $u(t)$ qui se caractérise par une pointe u de la commande sur un intervalle infiniment petit s'appelle « petite » variation.

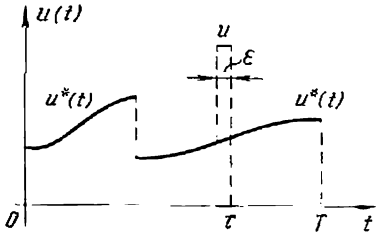


Fig. 2.12.

Notons que lors d'une « petite » variation l'accroissement $(u - u^*)$ ne doit nullement être aussi petit que l'on veut. Au contraire, les grandeurs u et u^* peuvent être quelconques, pourvu qu'elles ne dépassent pas les limites admissibles. Par exemple, si la commande est restreinte par la condition

$$|u| \leqslant U = \text{const}, \tag{2.243}$$

alors u, de même que u^*, doit satisfaire à cette condition, le module de leur différence pouvant varier entre 0 et $2U$.

L'idée d'une « petite » variation présente cet intérêt que, malgré la valeur *finie* de la différence $(u - u^*)$, l'influence de cette variation sur le mouvement ultérieur de l'objet est *infiniment petite*. De simples considérations physiques rendent évidente cette particu-

larité. En effet, on sait que l'action exercée sur un système par une courte impulsion quelconque est appréciée d'après la valeur de l'aire sous cette impulsion. L'aire $(u - u^*)\,\varepsilon$ de l'accroissement étant infiniment petite, son influence sur la trajectoire $\tilde{x}(t)$ $(t > \tau)$ est également infiniment petite. Dans ce qui suit nous justifions cette proposition.

Notons également que la « petite » variation diffère de la variation utilisée dans le calcul des variations classique. La variation classique doit être une fonction suffisamment lisse, alors que la « petite » variation représente un échelon. Ce n'est qu'en examinant la classe des fonctions $u(t)$ continues par morceaux qu'on peut faire intervenir la « petite » variation (fig. 2.12). Or c'est précisément cette dernière qu'on prend comme point de départ pour déduire le principe du maximum. Ainsi, la discontinuité, qui dans le calcul des variations classique est une pierre d'achoppement, devient un précieux outil dans la théorie du principe du maximum *).

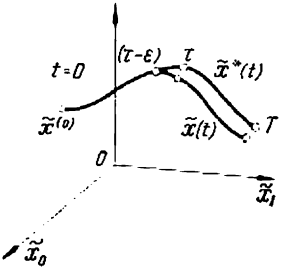


Fig. 2.13.

Après une variation de la commande sur l'intervalle infiniment petit $\tau - \varepsilon < t < \tau$ la trajectoire ultérieure $\tilde{x}(t)$, $t > \tau$, diffère de la trajectoire optimale $\tilde{x}^*(t)$. La différence entre ces grandeurs à l'instant $t = \tau$ est égale, aux infiniment petits d'ordre supérieur près, au produit de la différence des vitesses des variations $\left(\frac{d\tilde{x}}{dt} - \frac{d\tilde{x}^*}{dt}\right)_{t=\tau}$ par l'intervalle de temps ε:

$$\tilde{x}(\tau) - \tilde{x}^*(\tau) = \varepsilon \left\{ \left(\frac{d\tilde{x}}{dt}\right)_{t=\tau} - \left(\frac{d\tilde{x}^*}{dt}\right)_{t=\tau} \right\} =$$

$$= \varepsilon \{ \tilde{f}[\tilde{x}(\tau),\ u(\tau)] - \tilde{f}[\tilde{x}(t),\ u^*(\tau)] \}. \qquad (2.244)$$

Cette différence étant infiniment petite, mais différente du zéro, il existe, en général, un certain écart entre les trajectoires $\tilde{x}(t)$ et $\tilde{x}^*(t)$ pour $t > \tau$. En effet, bien que pour $t > \tau$ $u^*(t)$ soit la même pour les deux trajectoires, d'après (2.244) à l'instant $t = \tau$ les valeurs de $\tilde{x}(\tau)$ et de $\tilde{x}^*(\tau)$, qui sont des « conditions initiales » pour l'intervalle $\tau < t < T$, diffèrent l'une de l'autre. Par conséquent, pour $t > \tau$, $\tilde{x}(t)$ et $\tilde{x}^*(t)$ ne coïncident pas. Cependant, la différence $\tilde{x}(\tau) - \tilde{x}^*(\tau)$ étant infiniment petite, la trajectoire ultérieure $\tilde{x}(t)$ différera *d'une valeur infiniment petite* de $\tilde{x}^*(t)$. La figu-

*) Il est à noter que la « petite » variation a été utilisée également dans d'autres généralisations des méthodes variationnelles classiques.

re 2.13 illustre bien cette divergence des trajectoires $\tilde{x}(t)$ et $\tilde{x}^*(t)$ pour $t > \tau$.

Introduisons le vecteur de la variation de la trajectoire $\delta\tilde{x}(t)$ de coordonnées $\delta\tilde{x}_j$ $(j = 0, \ldots, n+1)$, défini par l'égalité

$$\delta\tilde{x}(t) = \tilde{x}(t) - \tilde{x}^*(t) \quad (\tau \leqslant t < T). \tag{2.245}$$

En vertu de (2.244) la valeur « initiale » de la variation pour $t = \tau$ est égale à

$$\delta\tilde{x}(\tau) = \varepsilon\{\tilde{f}[\tilde{x}(\tau), \quad u(\tau)] - \tilde{f}[\tilde{x}(\tau), \quad u^*(\tau)]\}. \tag{2.246}$$

Etant donné que pour $t > \tau$ $\tilde{x}(t)$ diffère de $\tilde{x}^*(t)$ aussi peu que l'on veut, la variation $\delta\tilde{x}(t)$ de la trajectoire est infiniment petite. Sa loi peut être donc établie à partir des équations linéaires pour faibles variations de $\tilde{x}(t)$ qu'on appelle *équations aux variations*. Les équations aux variations s'obtiennent à partir des équations principales

$$\frac{d\tilde{x}_j}{dt} = \tilde{f}_j(\tilde{x}_0, \tilde{x}_1, \ldots, \tilde{x}_{n+1}, u) \quad (j = 0, \ldots, n+1), \tag{2.247}$$

en remplaçant les $\tilde{x}_j$ par $\tilde{x}_j + \delta\tilde{x}_j$ $(j = 0, \ldots, n+1)$ et en rejetant dans le développement en série de $\tilde{f}_j$ par rapport à $\delta\tilde{x}_i$ les termes d'ordres de petitesse supérieurs, on a

$$\frac{d(\tilde{x}_j + \delta\tilde{x}_j)}{dt} = \tilde{f}_j(\tilde{x}_0 + \delta\tilde{x}_0, \ldots, \tilde{x}_i + \delta\tilde{x}_i, \ldots, \tilde{x}_{n+1} + \delta\tilde{x}_{n+1}, u) =$$
$$= \tilde{f}_j(\tilde{x}_0, \ldots, \tilde{x}_i, \ldots, \tilde{x}_{n+1}, u) +$$
$$+ \sum_{i=0}^{n+1} \delta\tilde{x}_i \frac{\partial\tilde{f}_j}{\partial\tilde{x}_i}(\tilde{x}_0, \ldots, \tilde{x}_i, \ldots, \tilde{x}_{n+1}, u) + o(\delta\tilde{x}). \tag{2.248}$$

En rejetant le terme $o(\delta\tilde{x})$ qui contient des termes d'ordre de petitesse supérieur au premier et compte tenu de (2.247), on aboutit aux équations aux variations linéaires pour $\delta\tilde{x}_i$:

$$\frac{d(\delta\tilde{x}_j)}{dt} = \sum_{i=0}^{n+1} \delta\tilde{x}_i \frac{\delta\tilde{f}_j(\tilde{x}, u)}{\delta\tilde{x}_i} \quad (j = 0, \ldots, n+1). \tag{2.249}$$

Ces équations peuvent être intégrées sous les conditions initiales (2.246). Notre intérêt est porté tout d'abord à la grandeur $(\delta\tilde{x})_{t=T}$, plus spécialement à la valeur de la coordonnée $\delta\tilde{x}_0$ à l'instant $t = T$. En effet, d'après (2.239) cette valeur est une variation δQ du critère Q dans l'intervalle $\tau - \varepsilon < t < \tau$ due à une « petite » variation. Comme la commande optimale $u^*(\tau)$ assure la valeur de Q minimale, toute autre commande $u(\tau)$ ne peut qu'augmenter la valeur

de Q. Par conséquent

$$\delta Q = (\delta \tilde{x}_0)_{t=T} \geqslant 0. \tag{2.250}$$

Cette relation peut être mise sous la forme

$$-\delta Q = -(\delta \tilde{x}_0)_{t=T} = \langle \delta \tilde{x}(T), \tilde{\psi}(T) \rangle \leqslant 0, \tag{2.251}$$

où $\tilde{\psi}(T)$ est un vecteur choisi de façon que le produit scalaire de $\delta \tilde{x}(T)$ par $\tilde{\psi}(T)$ soit égal à $\delta \tilde{x}_0(T)$. Il est clair que la coordonnée de ce vecteur $\tilde{\psi}_0(T) = -1$, alors que les autres coordonnées $\tilde{\psi}_j(T) = 0$ $(j = 1, \ldots, n+1)$. Ainsi

$$\tilde{\psi}(T) = (1, 0, \ldots, 0). \tag{2.252}$$

Le produit scalaire (2.251) représente l'effet que la « petite » variation, apparue à l'instant $t=\tau$, produit sur le but final, c'est-à-dire sur le critère d'optimalité Q ou sur $\tilde{x}_0$, à l'instant $t = T$. Soulignons que l'idée maîtresse de la déduction du principe du maximum consiste en ce que cette influence peut être évaluée à l'aide des équations *linéaires* aux variations (2.249), car elle est aussi petite que l'on veut. La linéarité des équations simplifie beaucoup la démonstration. Ainsi, deux « petites » variations se manifestant à des intervalles de temps infiniment petits différents peuvent être considérées indépendamment l'une de l'autre par suite des propriétés d'additivité propres aux équations linéaires. Le choix de la valeur optimale de $u(\tau)$ à un instant quelconque peut donc être guidé *formellement* par le souci de rendre la valeur de l'accroissement $-\delta Q$ correspondant la plus grande possible, *indépendamment du processus de commande* à d'autres instants. Pour les commandes non optimales, cette grandeur, égale, d'après (2.251), au produit scalaire

$$-\delta Q = \langle \delta \tilde{x}(T), \tilde{\psi}(T) \rangle \tag{2.253}$$

sera négative. Elle ne devient nulle que pour $u(\tau) = u^*(\tau)$ en atteignant alors sa valeur maximale. Au fond, la condition (2.251) signifie qu'une commande non optimale est « pire » qu'une commande optimale, car l'effet qu'elle produit est plus faible.

L'expression (2.253) n'est pas assez commode, car pour la calculer il faut au préalable intégrer l'expression (2.249) et obtenir $\delta \tilde{x}(T)$ en fonction de la condition « initiale » $\delta \tilde{x}(\tau)$. Or (2.246) montre que c'est précisément $\delta \tilde{x}(\tau)$ qui est liée directement à $u(\tau)$. Donnons-nous pour tâche de rechercher le vecteur $\tilde{\psi}(t)$ tel qu'il vérifie la condition

$$\langle \delta \tilde{x}(t), \tilde{\psi}(t) \rangle = \langle \delta \tilde{x}(T), \tilde{\psi}(T) \rangle \quad (\tau \leqslant t \leqslant T). \tag{2.254}$$

Dans le cas particulier pour $t = \tau$ on arrive à l'égalité

$$\langle \delta\tilde{x}(\tau), \tilde{\psi}(\tau)\rangle = \langle \delta\tilde{x}(T), \tilde{\psi}(T)\rangle \tag{2.255}$$

et on peut juger de δQ d'après le premier membre de cette égalité, directement lié à la valeur cherchée de $u(\tau)$.

Il s'avère qu'on peut trouver une équation différentielle à laquelle satisfait le vecteur $\tilde{\psi}(t)$. En vertu de (2.254)

$$\langle \delta\tilde{x}(t), \tilde{\psi}(t)\rangle = \text{const} \quad (\tau \leqslant t \leqslant T). \tag{2.256}$$

On en tire

$$\frac{d}{dt}\langle \delta\tilde{x}(t), \tilde{\psi}(t)\rangle = 0 \quad (\tau \leqslant t \leqslant T), \tag{2.257}$$

ou

$$\left\langle \frac{d[\delta\tilde{x}(t)]}{dt}, \tilde{\psi}(t)\right\rangle + \left\langle \delta\tilde{x}(t), \frac{d\tilde{\psi}(t)}{dt}\right\rangle = 0 \quad (\tau \leqslant t \leqslant T). \tag{2.258}$$

Ecrivons cette égalité sous une forme développée

$$\sum_{j=0}^{n+1} \frac{d[\delta\tilde{x}_j(t)]}{dt} \cdot \tilde{\psi}_j(t) + \sum_{i=0}^{n+1} \delta\tilde{x}_j(t) \frac{d\tilde{\psi}_i(t)}{dt} = 0. \tag{2.259}$$

Remplaçons $\frac{d(\delta\tilde{x}_j)}{dt}$ par ses valeurs tirées de (2.249), on a

$$\sum_{j=0}^{n+1} \tilde{\psi}_j(t) \sum_{i=0}^{n+1} \delta\tilde{x}_i \frac{\partial\tilde{f}_j(\tilde{x}, u)}{\partial\tilde{x}_j} + \sum_{i=0}^{n+1} \delta\tilde{x}_i(t) \frac{d\tilde{\psi}_i(t)}{dt} = 0. \tag{2.260}$$

En modifiant l'ordre de sommation par rapport à i et j dans le premier terme, on aboutit à l'expression

$$\sum_{i=0}^{n+1} \delta\tilde{x}_i \left\{ \sum_{j=0}^{n+1} \tilde{\psi}_j(t) \frac{\partial\tilde{f}_j(\tilde{x}, u)}{\partial\tilde{x}_i} + \frac{d\tilde{\psi}_i(t)}{dt} \right\} = 0. \tag{2.261}$$

D'après (2.257), le premier membre de (2.261) est identiquement nul *quels que soient* $\delta\tilde{x}_j$. La condition nécessaire et suffisante en est l'égalité à zéro de l'expression entre accolades de (2.261), d'où l'on tire

$$\frac{d\tilde{\psi}_i(t)}{dt} = -\sum_{j=0}^{n+1} \tilde{\psi}_j(t) \frac{\partial\tilde{f}_j(\tilde{x}, u)}{\partial\tilde{x}_i} \quad (i = 0, \ldots, n+1). \tag{2.262}$$

Les égalités (2.262) représentent un ensemble d'équations différentielles, linéaires par rapport à $\tilde{\psi}_j$. On voit sans peine que les équations (2.262) et (2.231) sont identiques. Par conséquent, en procédant comme indiqué, on aboutit également aux équations

adjointes du système principal (2.249). Ces équations doivent être résolues pour les conditions aux limites (2.252).

Examinons maintenant la grandeur [cf. (2.255) et (2.251)]

$$-\delta Q = \langle \delta \tilde{x}(\tau), \tilde{\psi}(\tau) \rangle \leqslant 0. \tag{2.263}$$

En y portant $\delta \tilde{x}(\tau)$ tirée de (2.246) on tombe après simplification par ε sur

$$\langle \tilde{f}[\tilde{x}(\tau), u(\tau)], \tilde{\psi}(\tau) \rangle - \langle \tilde{f}[\tilde{x}(\tau), u^*(\tau)], \tilde{\psi}(\tau) \rangle \leqslant 0. \tag{2.264}$$

Introduisons maintenant la grandeur

$$\tilde{H} = \langle \tilde{f}[\tilde{x}(\tau), u(\tau)], \tilde{\psi}(\tau) \rangle. \tag{2.265}$$

L'inégalité (2.264) montre que le maximum de $\tilde{H}$ est obtenu pour la commande optimale $u^*(\tau)$. On en déduit le principe du maximum : $u(\tau)$ doit être choisie de façon à maximiser la valeur de $\tilde{H}$. Tous ces raisonnements se généralisent aisément au cas de $r > 1$, en remplaçant u par $\bar{u}$.

Considérons maintenant le mode d'application du principe du maximum. Pour rechercher la trajectoire optimale $\tilde{x}^*(t)$ et la commande optimale $\bar{u}^*(t)$ on résout simultanément deux systèmes d'équations, le système principal et adjoint. Supposons que le système démarre à $t = 0$ à partir d'un certain point $\tilde{x} = \tilde{x}^{(0)}$. Imposons-nous également une certaine valeur initiale $\tilde{\psi}^{(0)}$ du vecteur $\tilde{\psi}$ qui nous est inconnue d'avance. La valeur du vecteur $\bar{u}$ au point initial $\tilde{x}^{(0)}$ est choisie à partir de la condition $\tilde{H} = \max$ de façon qu'elle maximise le produit scalaire

$$(\tilde{H})_{\tilde{x}=\tilde{x}^{(0)}} = \langle \tilde{f}^{(0)} \tilde{\psi}^{(0)} \rangle = \left\langle \frac{d\tilde{x}[\tilde{x}^{(0)}, \bar{u}]}{dt}, \tilde{\psi}^{(0)} \right\rangle. \tag{2.266}$$

Après avoir choisi la commande $(\bar{u})_{t=0}$ on peut calculer à partir des systèmes d'équations principal (2.235) et adjoint (2.233) les accroissements $\Delta\tilde{\psi}_i$ et $\Delta\tilde{x}_i$ sur un intervalle de temps Δt suffisamment petit. Par conséquent on peut trouver les valeurs des vecteurs $\tilde{\psi}$ et $\tilde{x}$ en un nouveau point de la trajectoire optimale, voisin du point initial et correspondant à l'instant $t = \Delta t$. Toute la procédure décrite est reprise pour le nouveau point : on recherche une nouvelle valeur optimale $\bar{u}^*$ et de nouveaux accroissements $\Delta\tilde{\psi}_i$ et $\Delta\tilde{x}_i$, le point représentatif se déplace en position successive correspondant à l'instant le plus proche $t = 2\Delta t$, etc. En procédant ainsi de proche en proche on trace toute la trajectoire optimale.

La procédure exposée de la construction d'une trajectoire optimale compte jusqu'à présent un point obscur. On ne sait comment choi-

sir les valeurs initiales des coordonnées $\tilde{\psi}_i^{(0)}$ du vecteur $\tilde{\psi}$. Il se trouve que le choix de ces valeurs est lié aux conditions aux limites du problème.

Considérons quelques cas particuliers.

1. Problème à extrémité libre de la trajectoire et à temps fixé T.

Le problème qui consiste à minimiser la fonctionnelle Q se ramène, comme nous l'avons montré dans ce qui précède, à minimiser la valeur $\tilde{x}_0(T)$ en introduisant la variable $\tilde{x}_0$. On sait de ce qui précède qu'il faut choisir à cet effet la commande $\bar{u}$ telle que la projection du vecteur $\frac{d\tilde{x}}{dt}$ sur la direction $\tilde{\psi}$ soit maximale. Pour le «dernier» intervalle de temps infiniment petit $T - \Delta t < t < T$ la direction correspondante $\tilde{\psi}$ est donnée par la formule (2.252), dont le sens est facile à expliquer. Pour minimiser au possible l'accroissement $\tilde{x}_0(T)$ pour $t = T - \Delta t$ il faut orienter le vecteur $\tilde{\psi}$ dans une direction inverse à $\tilde{x}_0$. Mais, comme on le voit sans peine, cette direction $\tilde{\psi}$ coïncide avec celle du vecteur $\tilde{\psi}(T)$ définie par la formule (2.252).

Ainsi, dans le problème considéré il faut trouver une solution $\tilde{\psi}(t)$ telle qu'elle vérifie les conditions finales

$$\tilde{\psi}_0^{(T)} = -1, \quad \tilde{\psi}_i^{(T)} = 0 \quad (i = 1, \ldots, n+1). \tag{2.267}$$

Les valeurs initiales $\tilde{\psi}_i^{(0)}$ doivent être choisies de façon que les valeurs finales $\tilde{\psi}_i^{(T)}$ soient égales aux grandeurs imposées et définies par (2.267).

Ainsi, les solutions des systèmes d'équations principal et adjoint doivent satisfaire aux conditions *initiales* $\tilde{x}^{(0)}$ pour le vecteur $\tilde{x}$ et aux conditions *finales* (2.267) pour le vecteur $\tilde{\psi}$. En général, il faut résoudre le problème d'intégration d'un système d'équations par rapport à $\tilde{x}$ et à $\tilde{\psi}$ pour les conditions *aux limites* données. Les $n+2$ conditions initiales pour $\tilde{x}^{(0)}$ et les $n+2$ conditions finales (2.264) pour $\tilde{\psi}^{(T)}$ donnent toutes les conditions aux limites nécessaires pour la résolution du problème.

2. Problème de la commande en temps minimal à extrémité de trajectoire $\bar{x}^{(T)}$ fixée dans un espace de phase x et à temps T non fixé d'avance.

On demande de choisir une commande telle que le point représentatif $\bar{x}$ soit déplacé de la position initiale $\bar{x}^{(0)}$ en position finale indiquée $\bar{x}^{(T)}$ en un temps minimal possible (fig. 2.14).

Pour résoudre conjointement les systèmes d'équations principal et adjoint (2.236) et (2.237) tout en choisissant la commande $\bar{u}$ qui

maximise la grandeur H en chaque point de la trajectoire optimale, il faut connaître en tout $2n$ conditions initiales. Ce sont les valeurs initiales $x_1^{(0)}, \ldots, x_n^{(0)}$ des coordonnées du point $\bar{x}$ dans l'espace de phase et les valeurs initiales $\psi_1^{(0)}, \ldots, \psi_n^{(0)}$ du vecteur $\bar{\psi}$. Les premières n valeurs sont imposées. Quant aux n coordonnées initiales du vecteur $\bar{\psi}$, leur choix doit être tel qu'elles vérifient les conditions aux limites imposées à l'extrémité de la trajectoire optimale, c'est-à-dire les valeurs imposées $x_1^{(T)}, \ldots, x_n^{(T)}$ des coordonnées du point final $\bar{x}^{(T)}$. On obtient ainsi n conditions pour n inconnues $\psi_i^{(0)}$ $(i = 1, \ldots, n)$.

Il n'existe pas de règle générale pour le choix des valeurs initiales $\psi_i^{(0)}$ des coordonnées du vecteur $\bar{\psi}$. Choisissons au hasard le vecteur $\bar{\psi}^{(0)}$ (fig. 2.14). Construisons par le procédé exposé ci-dessus la trajectoire optimale M_0M_1. Elle ne passera pas en général par le point nécessaire M_T pour lequel $\bar{x} = \bar{x}^{(T)}$. Introduisons alors une autre valeur du vecteur $\bar{\psi}^{(0)}$ et construisons une autre trajectoire M_0M_2 et ainsi de suite, tant que cette trajectoire ne passe par le point M_T. Cette procédure peut être améliorée en introduisant une mesure de distance r dans l'espace de phase depuis la trajectoire optimale M_0M_1 jusqu'au point M_T concerné [3.25]. Cette mesure peut être donnée, par exemple, par la plus petite des distances euclidiennes des points de la courbe M_0M_1 à M_T. Alors, en procédant comme indiqué plus haut il faut choisir les coordonnées $\psi_1^{(0)}, \ldots, \psi_n^{(0)}$ du vecteur $\bar{\psi}^{(0)}$ telles qu'elles minimisent la grandeur r qui en dépend

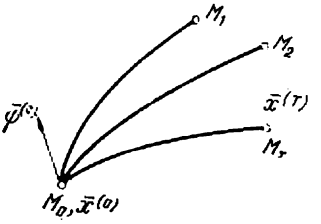


Fig. 2.14.

$$r = r(\psi_1^{(0)}, \ldots, \psi_n^{(0)}) = \min. \tag{2.268}$$

Ce minimum, certes, doit être égal à zéro. La fonction r peut donner lieu à plusieurs minimums mais la valeur recherchée constitue le minimum minimorum.

Ainsi, la procédure consiste en premier lieu à réaliser une série de maximisations de H par rapport à $\bar{u}$ pour tout petit intervalle de temps Δt d'où il résulte le « tracé » de la trajectoire optimale M_0M_j dans l'espace de phase. Pour toute trajectoire optimale construite de cette façon on calcule la valeur de r correspondante. Ensuite, par un choix convenable de $\psi_i^{(0)}$ $(i = 1, \ldots, n)$ on minimise r qui est leur fonction (2.268) en la réduisant à zéro. Ce n'est qu'alors que s'achève la résolution du problème. A l'heure actuelle il existe un

appareillage de synthèse automatique des systèmes optimaux [3.25, 3.71], [6.4, 6.6], qui réalise automatiquement cette procédure. Le principe qui régit cet appareillage est celui de l'exploration automatique. La résolution du problème par ces machines consiste en la recherche automatique « rapide » de la commande optimale $\bar{u}$ qui maximise la fonction $H(\bar{u})$ sur tout intervalle Δt. Il en résulte le « tracé » de la trajectoire optimale M_0M_i et la définition simultanée de sa distance r du point $\bar{x}^{(T)}$. Cette procédure s'accompagne de la recherche automatique « lente » des valeurs $\psi_i^{(0)}$ minimisant la fonction r.

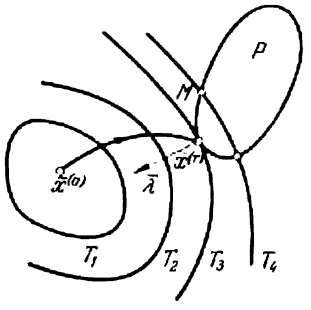


Fig. 2.15.

Pour les systèmes à objets linéaires il existe des méthodes permettant d'obtenir les valeurs $\psi_i^{(0)}$ par itération [2.39].

3. Problème de la commande en temps minimal à domaine P fini à n dimensions fixé dans l'espace de phase $\bar{x}$ et à temps T non fixé d'avance.

On demande de réaliser une commande telle que le point représentatif $\bar{x}$ se déplace de sa position initiale donnée $\bar{x}^{(0)}$ en un point $\bar{x}^{(T)}$ d'un certain domaine donné P à n dimensions de l'espace de phase en un temps minimal T. Ni le point $\bar{x}^{(T)}$ ni le temps T ne sont fixés d'avance.

Pour résoudre ce problème il faut connaître les n valeurs initiales $\psi_i^{(0)}$ $(i = 1, \ldots, n)$ du vecteur $\bar{\psi}$, puisque des $2n$ conditions initiales des systèmes d'équations principal et adjoint les n valeurs des coordonnées $x_1^{(0)}, \ldots, x_n^{(0)}$ sont données. Par conséquent, il reste à ajouter les n conditions aux limites. Ce sont ce qu'on appelle les *conditions de transversalité* imposées aux coordonnées du vecteur $\bar{\psi}$ au point final $\bar{x}^{(T)}$ de la trajectoire.

Pour obtenir les conditions de transversalité, considérons l'isosurface $T = \text{const}$ dans l'espace de phase (fig. 2.15). Convenons de donner ce nom au lieu géométrique des points de l'espace de phase qu'on peut atteindre en un temps T à partir du point initial $\bar{x}^{(0)}$ avec la commande optimale $\bar{u}^*$. Par exemple, en un temps T_1 on peut arriver du point $\bar{x}^{(0)}$ aux points de l'isosurface T_1, en un temps $T_2 > T_1$ aux points de l'isosurface T_2, etc. On peut démontrer que ces isosurfaces sont convexes. L'isosurface T_2 de la figure 2.15 ne coupe pas la surface délimitant le domaine P de l'espace de phase. Cela signifie que le temps T_2 est insuffisant pour atteindre le domaine P. En augmentant T on obtient de nouvelles isosurfaces englobant

celles à T moindre. En poursuivant ce processus de création de nouvelles isosurfaces $T = \text{const}$ qui englobent les points de l'espace de phase toujours de plus en plus éloignés du point initial, on aboutit finalement (si le problème a une solution) à l'isosurface $T = T_3$ *tangente* à la surface délimitant le domaine P. Soit $\overline{x}^{(T)}$ le point de tangence. C'est précisément le point final de la trajectoire optimale recherchée, matérialisée sur la figure 2.15 par un trait continu. En effet, dans les conditions d'une commande optimale, le point $\overline{x}^{(T)}$ peut être atteint en un temps $T = T_3$. Un autre point M du domaine P ne peut être atteint qu'en un temps $T_4 > T_3$, alors la trajectoire qui réunit $\overline{x}^{(0)}$ au point M, par exemple, ne sera pas optimale.

Soit l'équation de la surface délimitant le domaine P

$$\varphi(x_1, \ldots, x_n) = 0, \tag{2.269}$$

où φ est la fonction dérivable par rapport à tous les paramètres x_i $(i = 1, \ldots, n)$.

La construction de la figure 2.15 montre que la surface qui délimite le domaine P, si elle vérifie certaines conditions, et l'isosurface $T = T_3$ ont au point $\overline{x}^{(T)}$ un vecteur normal commun $\overline{\lambda}$.

On peut prendre alors comme normale $\overline{\lambda}$ le gradient φ, c'est-à-dire le vecteur

$$\overline{\lambda} = \operatorname{grad} \varphi = \left(\frac{\partial \varphi}{\partial x_1}, \ \ldots, \ \frac{\partial \varphi}{\partial x_n}\right). \tag{2.270}$$

Convenons que la surface $\varphi = 0$ est privée de points singuliers où tous les $\frac{d\varphi}{dx_i}$ deviennent simultanément nuls. Le vecteur grad φ est alors défini pour tout point de la surface (2.269).

La figure 2.15 montre également que la commande optimale intervenant dans le « dernier » intervalle de temps infiniment petit $T - \Delta t < t < T$, où $\Delta t \to 0$, consiste à déplacer le point représentatif le plus vite possible de l'isosurface $T_3 - \Delta t = \text{const}$ à l'isosurface $T_3 = \text{const}$. A cet effet il faut assurer une valeur maximale de la projection du vecteur $\frac{d\overline{x}}{dt}$ sur la direction du vecteur $\overline{\lambda}$. Par conséquent, on est dans le cas où la direction du vecteur $\overline{\psi}(T)$ coïncide avec celle de $\overline{\lambda}$ et on peut poser

$$\overline{\psi}(T) = \operatorname{grad} \varphi = \left(\frac{\partial \varphi}{\partial x_1}, \ \ldots, \ \frac{\partial \varphi}{\partial x_n}\right) \tag{2.271}$$

d'où l'on tire

$$\psi_i(T) = \frac{\partial \varphi}{\partial x_i} \quad (i = 1, \ldots, n). \tag{2.272}$$

Les conditions (2.271) sont les conditions de transversalité. Le temps T n'étant pas fixé, les valeurs $\psi_1^{(0)}, \ldots, \psi_n^{(0)}$ et la grandeur

T, inconnues d'avance, sont les $n+1$ inconnues qui nécessitent pour leur définition $n+1$ conditions. Ces conditions sont données par $n+1$ relations (2.269) et (2.272). En effet, si nous nous imposons les valeurs $\psi_i^{(0)}$ et T, les coordonnées $x_i(T)$ et $\psi_i(T)$ seront les fonctions de ces $n+1$ inconnues. En portant ces fonctions dans les $n+1$ relations (2.269) et (2.272) nous pouvons nous attendre que les $n+1$ équations obtenues aux $n+1$ inconnues $\psi_i^{(0)}$ et T aient une solution. Bien entendu, les difficultés de la résolution du problème augmentent par rapport aux cas précédents, par suite de la complexité de définition des conditions aux limites. Si le domaine P se contracte en un point, nous sommes ramenés au problème précédent. Dans ce cas les conditions de transversalité perdent leur sens, mais, en revanche, on voit apparaître les coordonnées du point final de la trajectoire, donc le nombre total d'équations définissant les valeurs inconnues $\psi_i^{(0)}$ devient de nouveau suffisant.

Les conditions du problème peuvent être généralisées si au lieu du point initial fixé $\bar{x}^{(0)}$ on considère un certain domaine initial P_0 des points de l'espace de phase d'où peut partir le point représentatif. Dans ce cas les conditions de transversalité analogues à celles décrites plus haut sont imposées également à la valeur initiale du vecteur $\bar{\psi}$. Pour plus de détails nous adressons le lecteur à la monographie de L. Pontriaguine et coll. [2.21] qui donne une démonstration rigoureuse et générale du principe du maximum compte tenu des conditions de transversalité.

4. Problème de réalisation d'une commande en temps minimal à domaine fini fixé à s dimensions ($s < n$).

Dans ce cas, plus général que le précédent, le point final $\bar{x}^{(T)}$ de la trajectoire doit reposer sur la variété M_s de dimension s, $1 < s < n$, les coordonnées des points $\bar{x}$ appartenant à M_s étant données par le système d'équations

$$\varphi_i(x_1, \ldots, x_n) = \varphi_i(\bar{x}) = 0 \quad (i = 1, \ldots, n-s). \tag{2.273}$$

Ainsi, deux équations du type (2.273) définissent dans un espace tridimensionnel une variété unidimensionnelle M_s, c'est-à-dire une certaine ligne.

Le vecteur $\bar{g}$ normal à la variété M_s peut être donné sous la forme suivante :

$$\bar{g} = \sum_{\beta=1}^{\beta=n-s} \lambda_\beta \operatorname{grad} \varphi_\beta(\bar{x}), \tag{2.274}$$

où $\bar{x}$ vérifie les équations (2.273) et λ_β sont certains nombres. De plus, nous considérons qu'en aucun point de la variété M_s aucun des vecteurs $\operatorname{grad} \varphi_\beta(\bar{x})$ ne devient nul.

Il est facile de vérifier que le vecteur $\bar{g}$ est orthogonal à tout vecteur infiniment petit $\delta\bar{x}$ appartenant à la variété M_s et issu du point $\bar{x}$. En effet, le vecteur $\delta\bar{x}$ repose sur toute surface $\varphi_\beta(\bar{x}) = 0$ (2.273) et, par conséquent, il est orthogonal à tout grad $\varphi_\beta(\bar{x})$. Le produit scalaire ci-dessous est égal à zéro

$$\langle \bar{g}, \delta\bar{x}\rangle = \sum_{\beta=1}^{\beta=n-s} \lambda_\beta \langle \operatorname{grad} \varphi_\beta(\bar{x}), \ \delta\bar{x}\rangle = 0, \tag{2.275}$$

c'est-à-dire les vecteurs $\bar{g}$ et $\delta\bar{x}$ sont orthogonaux.

Supposons d'abord que le temps n'est pas présent explicitement dans les équations de l'objet, on a

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}, \bar{u}), \tag{2.276}$$

où $\bar{x}$ est le vecteur à n dimensions et $\bar{u}$ le vecteur à r dimensions. La commande $\bar{u}$ cherchée doit minimiser l'intégrale

$$Q = \int_0^T G(\bar{x}, \bar{u})\, dt = \int_0^T f_0(\bar{x}, \bar{u})\, dt = x_0(T). \tag{2.277}$$

Le temps T peut ne pas être fixé. Dans un espace à n dimensions du vecteur $\bar{x}$ on peut imaginer des isosurfaces S_t = const, où

$$S_t = \int_0^t G(\bar{x}, \bar{u}^*)\, dt, \tag{2.278}$$

et $\bar{u}^*$ est une commande optimale. Les isosurfaces de ce type pour le cas particulier, où $G = 1$ et le temps T nécessaire est minimal, sont montrées sur la figure 2.15. Evidemment on peut construire de telles isosurfaces pour un cas plus général lorsque G n'est pas égal à 1. Le vecteur $\bar{\psi}$ est un gradient à l'isosurface S_t = const. Il en est ainsi parce que le vecteur $\bar{\psi}$ aux coordonnées $\left(-\frac{\partial S}{\partial x_i}\right)$ est le gradient à l'isosurface S = const, ce qui se déduit aisément de (2.210). Alors, en comptant le temps à rebours, nous pouvons établir l'identité des isosurfaces S = const de ce nouveau problème aux isosurfaces S_t = const [cf. (2.278)]. Les deux formules ne différant que par le signe, on a donc

$$\bar{\psi} = \operatorname{grad} S_t \tag{2.279}$$

(sans le signe moins, comme il en était dans la formule pour S).

Imaginons, comme dans le cas précédent, une famille d'isosurfaces S_t en extension associées à des valeurs de t toujours plus grandes.

Si, pour une faible valeur de t, la surface correspondante n'a pas encore de points communs avec la variété M_s, avec l'augmentation de t il arrivera un instant t_1 (si le problème admet une solution) où l'isosurface $S_t = \text{const}$ aura au moins un point commun avec la variété M_s (on admet que de tels points n'existaient pas pour $t < t_1$). Comme la tangence de l'isosurface et de la variété entraîne la colinearité de leurs vecteurs normaux $\overline{\psi}$ et $\overline{g}$, on tombe sur la condition de transversalité sous la forme *)

$$\overline{\psi}(t_1) = \overline{g} = \sum_{\beta=1}^{n-s} \lambda_\beta \operatorname{grad} \varphi_\beta [x(t_1)]. \tag{2.280}$$

Autrement dit, le point final $\overline{x}(t_1)$ vérifiant les conditions (2.273) doit donner lieu à des nombres λ_β, non nuls simultanément, vérifiant l'égalité (2.280).

L'égalité vectorielle (2.280) est équivalente à n égalités scalaires. Si l'on y ajoute encore $n - s$ conditions (2.273), on obtient en tout $2n - s$ équations aux $2n - s$ inconnues, c'est-à-dire $n - s$ constantes λ_β et n inconnues $\psi_i(t_1)$.

Si dans les équations (2.276) t est présent sous une forme explicite, la substitution $x_{n+1} = t$ ramène le problème au cas précédent à cette différence que les constructions se font dans un espace à $n + 1$ dimensions. De plus, dans ce nouvel espace, la variété M_{s+1}, équivalente à l'ancienne variété M_s de l'espace $\overline{x}$ à n dimensions, sera un cylindre ayant pour base la variété M_s de l'ancien espace et dont les génératrices sont des droites parallèles à l'axe des coordonnées $x_{n+1} \equiv t$. La forme des conditions de transversalité est ici la même que précédemment. Ayant fixé à T le temps de passage du point initial à la variété M_s, on ajoute aux équations de la variété une équation de plus

$$\varphi_{n-s+1}(\widetilde{x}) = x_{n+1} - T = 0 \tag{2.281}$$

qui doit être utilisee pour déterminer les conditions de transversalité.

Donnons à titre d'illustration de ce qui vient d'être dit un exemple de l'application du principe du maximum à un problème à extrémité libre de la trajectoire et à temps T fixé [2.17]. Proposons-nous de trouver une commande qui minimise l'intégrale

$$Q = \frac{1}{2} \int_0^T (x^2 + u^2)\, dt, \tag{2.282}$$

*) La linéarité des équations adjointes fait qu'il suffit de définir le vecteur $\overline{\psi}$ à un facteur constant près. En multipliant alors les conditions finales par un facteur constant quelconque on obtient une nouvelle valeur de $\overline{\psi}(t)$ qui constitue également une solution.

l'objet étant décrit par l'équation du premier ordre suivante:

$$\frac{dx}{dt} = -ax + u. \tag{2.283}$$

Soit la valeur $(x)_{t=0} = x^{(0)}$. Supposons également que la commande u ne subisse aucune restriction supplémentaire.

Introduisons les variables

$$\left.\begin{aligned} \tilde{x}_1(t) &= x(t), \\ \tilde{x}_0(t) &= \frac{1}{2}\int_2^1 (\tilde{x}_1^2 + u^2)\,dt, \end{aligned}\right\} \tag{2.284}$$

avec, de plus, $(\tilde{x}_0)_{t=0} = 0$. Le système d'équations s'écrit alors

$$\left.\begin{aligned} \frac{d\tilde{x}_0}{dt} &= \frac{1}{2}\tilde{x}_1^2 + \frac{1}{2}u^2 = \tilde{f}_0, \\ \frac{d\tilde{x}_1}{dt} &= -a\tilde{x}_1 + u = \tilde{f}_1. \end{aligned}\right\} \tag{2.285}$$

Composons maintenant la fonction $\tilde{H}$ d'après l'expression (2.228):

$$\tilde{H} = \tilde{\psi}_0 \frac{d\tilde{x}_0}{dt} + \tilde{\psi}_1 \frac{d\tilde{x}_1}{dt} = \tilde{\psi}_0 \frac{1}{2}(\tilde{x}_1^2 + u^2) + \tilde{\psi}_1(-a\tilde{x}_1 + u). \tag{2.286}$$

Ecrivons les équations adjointes pour $\tilde{\psi}_0$ et $\tilde{\psi}_1$ conformément à (2.227) et (2.231). On a

$$\left.\begin{aligned} &\frac{d\tilde{\psi}_0}{dt} = 0, \\ &\frac{d\tilde{\psi}_1}{dt} = -\sum_{j=0}^{1} \tilde{\psi}_j \frac{\partial \tilde{f}_j}{\partial \tilde{x}_1} = -\tilde{\psi}_0 \frac{\partial \tilde{f}_0}{\partial \tilde{x}_1} - \tilde{\psi}_1 \frac{\partial \tilde{f}_1}{\partial \tilde{x}_1} = -\tilde{\psi}_0 \tilde{x}_1 + \tilde{\psi}_1 a. \end{aligned}\right\} \tag{2.287}$$

D'après (2.252) les valeurs finales $\tilde{\psi}_i$ sont

$$\left.\begin{aligned} \tilde{\psi}_0(T) &= -1, \\ \tilde{\psi}_1(T) &= 0. \end{aligned}\right\} \tag{2.288}$$

Les conditions initiales imposées à $\tilde{x}$

$$\left.\begin{aligned} \tilde{x}_0^{(0)} &= 0, \\ \tilde{x}_1^{(0)} &= x^{(0)}. \end{aligned}\right\} \tag{2.289}$$

La commande u à choisir à chaque instant doit être telle qu'elle maximise $\tilde{H}$. En vertu des premières conditions de (2.287) et (2.288)

on a

$$\tilde{\psi}_0(t) = -1 = \text{const.} \tag{2.290}$$

Alors (2.286) donne

$$\tilde{H} = -\frac{\tilde{x}_1^2}{2} - \frac{u^2}{2} - a\tilde{\psi}_1\tilde{x}_1 + \tilde{\psi}_1 u. \tag{2.291}$$

En égalant la dérivée $\frac{d\tilde{H}}{du}$ à zéro on obtient la valeur optimale

$$u^* = \tilde{\psi}_1. \tag{2.292}$$

Portons cette valeur dans les équations pour $\tilde{x}_1$ et $\tilde{\psi}_1$ (la fonction $\tilde{x}_0$ ne faisant pas partie de ces équations ne présente aucun intérêt pour nous), il vient

$$\left.\begin{aligned} \frac{d\tilde{x}_1}{dt} &= -a\tilde{x}_1 + \tilde{\psi}_1, \\ \frac{d\tilde{\psi}_1}{dt} &= a\tilde{\psi}_1 + \tilde{x}_1. \end{aligned}\right\} \tag{2.293}$$

D'après (2.288) et (2.289) les conditions aux limites de ce système sont

$$\left.\begin{aligned} \tilde{x}_1(0) &= x^{(0)}, \\ \tilde{\psi}_1(T) &= 0. \end{aligned}\right\} \tag{2.294}$$

Les équations linéaires (2.293) se prêtent aisément à l'intégration. On a après intégration

$$\left.\begin{aligned} \tilde{x}_1(t) &= C_1 e^{pt} + C_2 e^{-pt}, \\ \tilde{\psi}_1(t) &= D_1 e^{pt} = D_2 e^{-pt}, \end{aligned}\right\} \tag{2.295}$$

où

$$p = \sqrt{a^2+1} \tag{2.296}$$

est la racine de l'équation caractéristique.

Les conditions (2.294) prennent la forme

$$\left.\begin{aligned} x^{(0)} &= C_1 + C_2, \\ 0 &= \tilde{\psi}_1(T) = D_1 e^{pT} + D_2 e^{-pT}. \end{aligned}\right\} \tag{2.297}$$

L'expression (2.293) pour $t = 0$ entraîne

$$\left.\begin{aligned} \left[\frac{d\tilde{x}_1}{dt}\right]_{t=0} &= C_1 p - C_2 p = -a\tilde{x}_1(0) + \tilde{\psi}_1(0) = -ax^{(0)} + D_1 + D_2, \\ \left[\frac{d\tilde{\psi}_1}{dt}\right]_{t=0} &= D_1 p - D_2 p = a\tilde{\psi}_1(0) + \tilde{x}_1(0) = a(D_1 + D_2) + x^{(0)}. \end{aligned}\right\} \tag{2.298}$$

La résolution simultanée de (2.298) et de (2.297) permet de définir les constantes C_1, C_2, D_1, D_2. Nous tirons notamment

$$D_1 = -\frac{x^{(0)}}{(p-a)+(p+a)e^{2pT}}, \quad D_2 = -\frac{x^{(0)}\,e^{2pT}}{(p-a)+(p+a)\,e^{2pT}}, \tag{2.299}$$

ce qui nous donne avec (2.292) et (2.295)

$$u^*(t) = \widetilde{\psi}_1(t) = \frac{x^{(0)}\,[e^{pt} - e^{2pt}\cdot e^{-pT}]}{(p-a)+(p+a)\,e^{2pT}}. \tag{2.300}$$

L'expression (2.300) donne la loi de la commande optimale en fonction du temps. Ayant trouvé $\widetilde{x}_1(t)$ et éliminé le temps t de $u^*(t)$ et de $\widetilde{x}_1(t)$, on aboutit à $u^* = u^*(\widetilde{x}_1)$.

Le principe du maximum s'étend naturellement aux systèmes continus à état discret à objets *linéaires* [2.17]. En effet, pour ces objets, même dans le cas des écarts importants, les équations des écarts sont également linéaires. Par conséquent, les raisonnements qui ont servi à la discussion du principe du maximum dans le cas des systèmes continus peuvent être repris, avec quelques modifications négligeables, pour l'étude des systèmes continus à état discret à objets linéaires. Toutefois, dans le cas général des systèmes continus à état discret non linéaires, une telle démonstration est impossible. En effet, les intervalles de temps entre les échantillons ainsi que le nombre total des valeurs discrètes sont des grandeurs finies. L'influence de la variation finale d'un échantillon $u(k)$, analogue de la « petite » variation pour le cas considéré, sur le but final, c'est-à-dire sur la valeur de Q, est également une valeur finie et non pas infiniment petite comme dans le cas des systèmes continus. Mais alors tout l'édifice des raisonnements qui précèdent s'effondre du fait qu'il est désormais impossible de parler de *petits* écarts entre la trajectoire en variation $\overline{x}$ et la trajectoire optimale $\overline{x}^*$. Et ce n'est pas tout ; on peut même imaginer un exemple contradictoire pour lequel, dans le cas d'une commande optimale, le principe du maximum, formulé comme ci-dessus, n'est plus valable.

Ces derniers temps il a été démontré (cf. [2.43]) que pour une certaine classe de systèmes continus à état discret non linéaires le principe du maximum a lieu. Dans le cas général ces systèmes répondent à une formulation bien plus « faible » des conditions nécessaires d'optimalité. Pour obtenir cette formulation « faible » il faut traiter une variation $\delta u(k)$ de la valeur discrète $u^*(k)$ telle que son influence sur la valeur finale de $Q = \overline{x}^{(T)}$ soit infiniment petite. Il est clair qu'en général ce n'est possible que lorsque la grandeur $\delta u(k)$ est elle-même infiniment petite. On peut reprendre alors les raisonnements ci-dessus et aboutir à un résultat d'aspect analogue. Or l'intérêt du résultat obtenu ne sera que *local* ; il n'est vrai que pour les variations suffisamment faibles $\delta u(k)$ et $\delta\widetilde{H}$ des grandeurs $u^*(k)$

et $\tilde{H}$. Dans ce cas on peut seulement démontrer que, si la commande u est optimale, elle rend stationnaire la grandeur $\tilde{H}$. Toute valeur stationnaire (maximum local, minimum local, valeur stationnaire propre) fait penser qu'on a là une commande optimale, mais on ignore toujours laquelle de ces valeurs constitue précisément la commande. On comprend aisément pourquoi dans sa formulation « faible » le principe du maximum présente peu d'intérêt.

Il nous a été impossible dans le cadre d'un chapitre de passer en revue toutes les méthodes mathématiques proposées et appliquées avec succès à la solution des problèmes de la théorie des systèmes optimaux. Citons, par exemple, les travaux de N. Krassovski [2.22, 2.23] qui ont à leur base les recherches de M. Kreine [2.24] relatives à l'analyse fonctionnelle, les ouvrages du théoricien polonais R. Kulikowski [2.25, 2.26] consacrés aussi à l'analyse fonctionnelle. Nous ne nous sommes pas attardés non plus sur ce qu'on appelle « régimes glissants ». Ces régimes donnent lieu à un nombre infiniment grand de sauts de la commande $\bar{u}(t)$. L'existence d'une commande optimale est traitée dans ce cas pour certaines classes de problèmes par F. Kirillova [2.27]. Pour les problèmes de ce type le principe du maximum a besoin d'être généralisé, ce qui a été réalisé par R. Gamkrélidzé [2.28]. Une autre approche de ces problèmes est proposée par V. Krotov [2.29].

CHAPITRE III

SYSTÈMES OPTIMAUX À INFORMATION COMPLÈTE SUR L'OBJET COMMANDÉ

§ 1. Problème de commande en temps minimal ; méthode de l'espace de phase

La théorie des systèmes à information complète sur l'objet commandé se développait d'abord comme une théorie de commande en temps minimal. Les systèmes utilisant le principe de commande en temps minimal sont devenus l'objet d'étude grâce à leur intérêt pratique. Déjà en 1935 D. Marianovski et D. Svetcharnik ont pris un brevet [3.1] pour un système de déplacement de cylindres de laminoir où été utilisée une réaction quadratique qui permettait d'obtenir une vitesse maximale. Plus tard la maison « Leeds and Northrup » (Etats-Unis) a appliqué un principe analogue à la construction du potentiomètre automatique « Speedomax ». En 1949 l'auteur du présent ouvrage a montré dans [3.2] que pour un objet linéaire de second ordre composé de deux circuits intégrateurs couplés en série le processus optimal d'asservissement comporte deux intervalles. Pour le premier la commande u, obéissant à la condition

$$|u| \leqslant U, \tag{3.1}$$

est maintenue à l'un de ses niveaux admissibles $\pm U$ et pour le deuxième, à l'autre niveau admissible. Dans cet ouvrage l'auteur a analysé les trajectoires optimales dans un plan de phase.

Dans l'ouvrage [3.3] publié en 1951, l'auteur a étudié, pour un objet linéaire de second ordre, les trajectoires optimales dans le plan de phase pour des conditions initiales différentes ; on y trouve également la comparaison des résultats théoriques et pratiques de simulation.

L'ouvrage [3.4] publié en 1952 généralise la formulation du problème d'asservissement aux systèmes d'ordre n pour un même type des conditions initiales ; son auteur présume que dans ce cas-là aussi les modules des grandeurs soumises aux restrictions doivent être maintenus à un niveau maximal admissible. Les processus décrits dans cet ouvrage ne sont pas tous des processus strictement optimaux, mais ils en sont proches. Ce même ouvrage donne l'énoncé du problème à plusieurs contraintes et émet l'hypothèse sur le type du processus optimal. hypothèse qui s'est avérée juste pour beaucoup d'autres problèmes.

Toute une série de travaux apparus simultanément avaient pour objet les différents systèmes optimaux de second ordre ou les principes particuliers dont l'application permettait d'améliorer les caractéristiques dynamiques du système sans le rendre encore optimal [3.25].

En 1953, dans [3.5] fut introduite la notion générale de processus optimal se déroulant dans un espace de phase de dimension n pour des conditions initiales quelconques et n'importe quelles actions extérieures admissibles. Dans cet ouvrage fut énoncé et démontré le théorème des n intervalles. Ce théorème a permis d'élaborer une méthode de synthèse d'une classe des systèmes optimaux d'ordre n (cf. [3.11]). Dans l'exposé du présent paragraphe l'auteur s'est basé sur les ouvrages [3.2, 3.5, 3.11].

Les résultats théoriques pour les systèmes de second ordre à racines conjuguées complexes ont été obtenus par le mathématicien américain D. W. Bushaw en 1953 [3.15, 3.17, 3.23].

Depuis 1954 la publication des ouvrages consacrés à la théorie des systèmes optimaux de commande en temps minimal s'est considérablement accrue [3.6-3.8, 3.12-3.14, 3.16, 3.18-3.24].

Donnons un exposé détaillé du problème de commande en temps minimal.

Soit l'objet B à commande continue dans un système à réaction caractérisé dans le cas général par l'équation vectorielle de la trajectoire

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}, \bar{u}, t), \tag{3.2}$$

où $\bar{x}$ est le vecteur de dimension n et le vecteur $\bar{u}$ a r coordonnées. La commande $\bar{u}$ est restreinte par la condition

$$\bar{u} \in \Omega(\bar{u}), \tag{3.3}$$

où $\Omega(\bar{u})$ est un certain domaine fermé admissible. L'expression (3.1) est un exemple concret illustrant la contrainte (3.3). Il importe de noter que l'extrémité du vecteur $\bar{u}$ peut reposer non seulement à l 'intérieur du domaine $\Omega(\bar{u})$ mais aussi sur sa frontière.

Nous allons poser que le processus idéal $\bar{x}(t)$ vérifie les égalités

$$x_i(t) = x_i^*(t) \quad (i = 1, 2, \ldots, n). \tag{3.4}$$

Ici $x_i(t)$ sont les coordonnées de l'objet et $x_i^*(t)$ les fonctions du temps données que nous allons considérer comme les coordonnées du vecteur $\bar{x}^*$:

$$\bar{x}^* = (x_1^*, x_2^*, \ldots, x_n^*). \tag{3.5}$$

Les fonctions $x_i^*(t)$ doivent également satisfaire à certaines contraintes. Pour expliquer leur sens, examinons l'espace de phase de dimen-

sion n du vecteur $\bar{x}$ (fig. 3.1.). Si l'état de l'objet B variait exactement d'après les conditions idéales (3.4), nous aurions l'égalité

$$\bar{x} = \bar{x}^*, \tag{3.6}$$

et la trajectoire de phase qui est l'hodographe du vecteur $\bar{x} = \bar{x}^*$ représenterait, par exemple, la trajectoire $N_0^* N^* N_T$. Soit N^* la position du point représentatif sur cette trajectoire à l'instant courant $t > 0$ et N_0^* la position de ce point à l'instant initial $t = 0$. Pourtant, l'état réel de l'objet à l'instant $t = 0$ est défini par le point représentatif N_0 qui ne coïncide pas avec N_0^*. L'état courant de l'objet à l'instant $t > 0$ est donné par le point N et la trajectoire de phase de l'objet est la courbe $N_0 N N_T$. Le but du système de régulation automatique consiste à faire coïncider en un temps minimal l'état réel N avec l'état imposé N^*. Supposons que les points N et N^* se confondent à l'instant $t = T$ en N_T. Supposons également qu'après cet instant, quand $t > T$, on peut choisir une commande $\bar{u}(t)$ telle qu'elle vérifie l'égalité (3.6). Toutefois, la marge de commande étant restreinte, les trajectoires $\bar{x}^*(t)$ ne sont pas toutes réalisables. Appelons *admissibles* les trajectoires réalisables $\bar{x} = \bar{x}^*(t)$. Il est clair que ces trajectoires sont les solutions de l'équation (3.2) sous la condition (3.3). Cette condition est précisément la contrainte à imposer aux trajectoires $\bar{x}^*(t)$. Demandons que le vecteur $\bar{u}$ se trouve à l'intérieur *) du domaine $\bar{\Omega}(u)$.

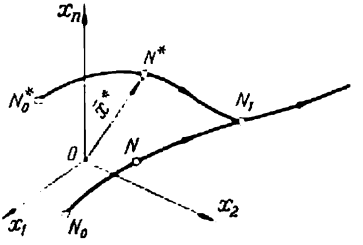


Fig. 3.1.

Le processus transitoire est celui de passage de l'état initial N_0 à l'état imposé N_T. La durée de ce processus est égale à T. Le système s'appelle optimal par rapport à la rapidité des processus transitoires si ces derniers donnent lieu à la condition

$$T = \min, \tag{3.7}$$

quelles que soient les conditions initiales et la fonction $\bar{x}^*(t)$ appartenant à la classe des fonctions admissibles ou bien à une sous-classe définie de cette classe. Dans ce dernier cas le système est dit en temps minimal pour la sous-classe indiquée des fonctions admissibles $\bar{x}^*(t)$. Le processus $\bar{x}(t)$ et la commande $\bar{u}(t)$ respective d'un tel système sont dits optimaux.

---

*) A l'intérieur, mais non pas sur la frontière, car le point N doit pouvoir « rattraper » le point N^* quelles que soient les conditions initiales.

Comme on le voit de la figure 3.1, le problème de recherche d'un processus optimal $\bar{x}(t)$ et d'une commande optimale $\bar{u}(t)$ peut être interprété comme un problème relatif à la rencontre de deux points N et N^* dans un espace de phase en temps minimal.

Il est souvent commode de remplacer l'espace de phase de $\bar{x}$ par un espace de phase de même dimension mais associé à l'erreur

$$\bar{\varepsilon} = \bar{x}^* - \bar{x} \tag{3.8}$$

de coordonnées $\varepsilon_i = x_i^* - x_i$ $(i = 1, 2, \ldots, n)$. A la fin du processus transitoire le vecteur de l'erreur devient nul. Par conséquent dans le nouvel espace le point représentatif du système, en partant d'une certaine position initiale, doit, en se déplaçant avec la plus grande vitesse admissible, rejoindre l'origine des coordonnées.

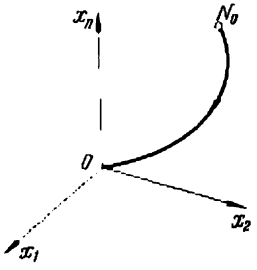


Fig. 3.2.

Le nouvel espace de phase s'obtient à partir de l'ancien si l'on reporte l'origine des coordonnées au point N^* en mouvement et constituant l'extrémité du vecteur $\bar{x}^*$ (fig. 3.1).

Remplaçons dans le nouvel espace de phase les notations ε_i $(i = 1, 2, \ldots, n)$ par les lettres x_i qui désignent désormais les coordonnées de l'erreur (fig. 3.2). Le problème de transfert du point représentatif du système d'une position initiale N_0 à l'origine des coordonnées O en un temps minimal ne diffère maintenant plus du problème de commande en temps minimal traité dans le chapitre précédent.

Pourquoi les points représentatifs N et N^* de la figure 3.1 ne peuvent-ils pas se rejoindre dans un laps de temps aussi petit que l'on veut? Ceci est impossible par suite des contraintes existant dans n'importe quel système réel, comme par exemple les contraintes (3.3). Ces contraintes rendent impossible le déplacement du point représentatif $\bar{x}$ dans l'espace de phase à des vitesses infiniment grandes.

Il arrive parfois que les contraintes sont imposées non seulement aux commandes $u_j(t)$ mais encore aux coordonnées $x_i(t)$ ou aux fonctions de ces coordonnées. Parfois le dernier type des contraintes peut amener à une contrainte appliquée à une commande non pas dans le schéma réel mais dans un certain autre schéma équivalent. Considérons à titre d'illustration l'exemple suivant.

Soit l'équation de la trajectoire de l'objet

$$b_0 \frac{d^2x}{dt^2} + b_1 \frac{dx}{dt} = u(t), \tag{3.9}$$

où b_i = const. C'est, par exemple, l'équation du mouvement d'un servomoteur à courant continu lorsque la grandeur d'entrée est la

tension recueillie sur l'induit et celle de sortie, l'angle de rotation de l'arbre. Supposons que la contrainte est imposée non pas à $u(t)$ mais à la dérivée seconde $\frac{d^2x}{dt^2}$:

$$\left|\frac{d^2x}{dt^2}\right| \leqslant M. \tag{3.10}$$

Dans l'exemple du servomoteur à courant continu la dérivée seconde est proportionnelle au courant de l'induit si le couple résistant est négligeable. Nous sommes alors dans le cas où le système (3.9) peut être remplacé par un système équivalent régi non par l'égalité (3.9) mais seulement par la contrainte (3.10)

$$\frac{d^2x}{dt^2} = v(t), \tag{3.11}$$

où $v(t)$ est la commande équivalente, ou la fonction de commande, qui vérifie la condition

$$|v(t)| \leqslant M. \tag{3.12}$$

Après avoir trouvé le processus optimal $x_{\text{opt}}(t)$ pour le système équivalent (3.11) on peut porter l'expression obtenue dans (3.9) et trouver la commande optimale réelle

$$u^*(t) = b_0 \frac{d^2x_{\text{opt}}}{dt^2} + b_1 \frac{dx_{\text{opt}}}{dt}. \tag{3.13}$$

Notons que dans ce cas le processus optimal se définit non par l'équation (3.9) de l'objet mais seulement par la contrainte (3.10). Toutefois, pour obtenir la commande optimale réelle $u^*(t)$ il faut connaître l'équation (3.9) de l'objet.

Un cas plus général des contraintes imposées aux coordonnées de l'objet et à la commande a été traité par R. Gamkrélidzé dans [2.18] (cf. également [2.45, 2.46]).

Rétrécissons maintenant la formulation du problème. Soit une seule commande $u(t)$ et soit l'objet à coordonnée de sortie x soumis à la contrainte du type

$$\left|a_0 \frac{d^nx}{dt^n} + a_1 \frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_nx\right| \leqslant M, \tag{3.14}$$

où $a_i = \text{const}$ et $a_0 > 0$. Si dans un cas particulier l'objet est défini par l'équation

$$a_0 \frac{d^nx}{dt^n} + a_1 \frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_nx = u(t), \tag{3.15}$$

la contrainte (3.14) se ramène à la condition

$$|u(t)| \leqslant M. \tag{3.16}$$

Pourtant, dans le cas plus général, comme nous l'avons dit plus haut, le premier membre de l'équation réelle de l'objet peut ne pas

coïncider avec l'expression (3.14). Dans tous ces cas seule la condition (3.14) suffit pour la construction du processus optimal.

Limitons notre exploration au cas où les racines de l'équation

$$a_0 p^n + a_1 p^{n-1} + \ldots + a_n = 0 \tag{3.17}$$

sont réelles et non positives. Autrement dit, les racines de (3.17) ne peuvent être que négatives et réelles ou nulles. Disons alors que l'équation (3.17) est caractéristique. Posons

$$a_0 \frac{d^n x}{dt^n} + a_1 \frac{d^{n-1} x}{dt^{n-1}} + \ldots + a_n x = v. \tag{3.18}$$

La contrainte (3.14) peut alors être mise sous la forme

$$|v| \leqslant M. \tag{3.19}$$

Dans le cas particulier d'une équation de l'objet du type (3.15), les expressions de $v(t)$ et de $u(t)$ coïncident. Toutefois, dans le cas plus général, comme nous l'avons dit plus haut, elles peuvent être différentes.

Pour les conditions indiquées le théorème des n intervalles est vrai. Voici son énoncé.

Le processus optimal $x(t)$ se répartit en n intervalles; dans chacun de ces intervalles il est décrit par l'équation

$$v = a_0 \frac{d^n x}{dt^n} + a_1 \frac{d^{n-1} x}{dt^{n-1}} + \ldots + a_n x = \sigma M, \tag{3.20}$$

où le nombre σ est constant sur chaque intervalle et

$$\sigma = \pm 1. \tag{3.21}$$

De plus, dans les intervalles adjacents, les signes de σ alternent.

Si, dans un cas particulier, $u = v$, le théorème des n intervalles peut être interprété comme suit: le système subit des commandes optimales $u(t)$ de type « en avant à toute vitesse » puis « en arrière à toute vitesse », au total n fois de suite.

Ce théorème a été démontré en 1953 dans [3.5] par une méthode élémentaire et presque sans aucun calcul. Pour mieux comprendre le principe de la démonstration examinons d'abord le cas particulier le plus simple traité en 1949 dans [3.2].

Soit, dans le cas particulier, l'équation de l'objet

$$a_0 \frac{d^2 x_1}{dt^2} = u, \tag{3.22}$$

la grandeur u étant soumise à la contrainte (3.16). Soient, ensuite, les conditions initiales de la forme

$$(x_1)_{t=0} = 0, \qquad \left(\frac{dx_1}{dt}\right)_{t=0} = 0. \tag{3.23}$$

Supposons qu'il faut faire passer l'objet en un temps minimal $T_{\min}$ à l'état

$$x_1 = x_{1f} = \text{const}, \quad \frac{dx_1}{dt} = 0. \tag{3.24}$$

Il se trouve qu'à cet effet il faut d'abord « accélérer » à une vitesse maximale avec $u = M$. Dans ces conditions, d'après (3.22), la fonction $\frac{dx_1}{dt}$ varie dans le temps d'après la loi linéaire

$$\frac{dx_1}{dt} = \int_0^t \frac{d^2x_1}{dt^2}\,dt = \int_0^t \frac{M}{a_0}\,dt = \frac{Mt}{a_0}. \tag{3.25}$$

La courbe *1* de la figure 3.3, *a* est celle de la vitesse $\frac{dx_1}{dt}$ du processus optimal. Au milieu de la trajectoire, lorsque $t = \frac{T}{2}$, l'« accélération » doit devenir « décélération » maximale, c'est-à-dire pendant le deuxième intervalle il faut maintenir $u = -M$. Ainsi la courbe optimale $\frac{dx_1}{dt}$ a une forme « triangulaire ». La distance parcourue x_1 étant définie par la formule

$$x_1 = \int_0^t \frac{dx_1}{dt}\,dt, \tag{3.26}$$

le processus optimal prend pour $t < T_{\min}$ la forme de deux segments de parabole (fig. 3.3,*b*) et pour $t \geqslant T_{\min}$ celle d'une droite horizontale $x_1 = x_{1f} = \text{const}$.

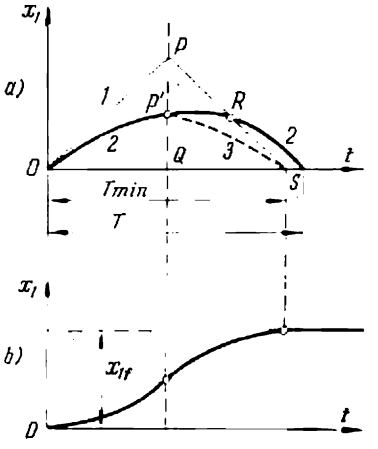


Fig. 3.3.

Montrons que le processus de la figure 3.3, *b* est optimal. Constatons en premier lieu que l'aire limitée par l'axe des abscisses et une courbe quelconque du processus transitoire $\frac{dx_1}{dt}$ doit être, dans le problème considéré, une grandeur constante égale à x_{1f}. En effet, si la durée du processus transitoire est T et si à sa fin l'égalité $x_1 = x_{1f}$ est vraie, (3.26) entraîne

$$\int_0^T \frac{dx_1}{dt}\,dt = x_{1f}. \tag{3.27}$$

Cette condition doit être vérifiée en particulier pour la courbe « triangulaire » *1* de la figure 3.3,*a*.

Considérons maintenant une autre courbe $\frac{dx_1}{dt}$, dans notre exemple la courbe *2* différente de la courbe « triangulaire » *1*. En vertu de la contrainte (3.16) et de l'équation (3.22) on a

$$\left|\frac{d^2x_1}{dt^2}\right| = \left|\frac{u}{a_0}\right| \leqslant \frac{M}{a_0} = \text{const}. \tag{3.28}$$

Alors, pour

$$0 \leqslant t \leqslant \frac{T_{\min}}{2}.$$

dans le premier intervalle la courbe *2* peut soit épouser la courbe *1*, soit passer au-dessous, mais ne peut pas se trouver au-dessus de cette courbe. En effet, d'après (3.28) la pente de la courbe *2* est inférieure ou égale à celle de la courbe *1*, alors que les valeurs initiales de ces courbes sont les mêmes. On en déduit que l'ordonnée du point P' est inférieure à celle du point P et l'intégrale $\int_0^{\frac{T_{\min}}{2}} \frac{dx_1}{dt}\,dt$ d'une courbe *2* quelconque est inférieure à celle de la courbe *1*. Or, x_{1f} étant la même, les intégrales (3.27) des deux courbes doivent être nécessairement égales. Par conséquent, pour $t > \frac{T_{\min}}{2}$, la courbe *2* ne peut pas passer comme indiqué en trait interrompu *3* de la figure 3.3,*a*. Elle coupe nécessairement la courbe *1* en un certain point R. La courbe *2* ne peut pas non plus couper une deuxième fois la courbe *1* car la pente de la courbe *2* est inférieure ou égale en valeur absolue à celle de la courbe *1*. Il en résulte que la courbe *2* rencontre l'axe des abscisses pour un $t = T$ plus grand que $T_{\min}$. On en déduit que dans le cas d'une courbe quelconque admissible, différente de *1*, la durée T du processus transitoire est plus grande que $T_{\min}$. Cela signifie précisément que la courbe *1* correspond bien au processus optimal. La durée $T_{\min}$ de ce processus se définit aisément à partir de la condition (3.27). Cette condition pour la courbe *1* traduit le fait que l'aire du triangle qu'elle forme, de base OS et de hauteur PQ, est égale à x_{1f}:

$$\frac{1}{2} PQ \cdot OS = \frac{1}{2}\left(\frac{M}{a_0}\frac{T_{\min}}{2}\right) \cdot T_{\min} = x_{1f}. \tag{3.29}$$

Il s'ensuit que

$$T_{\min} = 2\sqrt{\frac{a_0 x_{1f}}{M}}. \tag{3.30}$$

On voit de cette formule que la durée $T_{\min}$ d'un processus transitoire optimal est d'autant plus petite que la valeur maximale admissible M de la commande est plus grande. Pourtant la valeur de M étant finie, celle de $T_{\min}$ l'est également.

Dans le cas général, quelles que soient les conditions initiales pour x, la consigne admissible x^* et l'ordre n de l'équation des contraintes (3.18) [si $v = u$, cette dernière se ramène à l'équation de l'objet (3.15)], d'après le théorème des n intervalles, la fonction de commande $v(t)$ a la forme de la courbe représentée sur la figure 3.4. Dans notre exemple, dans le premier et le dernier intervalle la grandeur σ est égale à $+1$, c'est-à-dire $v = +M$. Tout le processus de variation de $v(t)$ est composé de n intervalles, en général, de durées différentes; par ailleurs, dans tout intervalle on a $\sigma = \text{const}$, tandis que dans les intervalles adjacents les signes de σ sont opposés. Cela signifie que la « marche en avant à toute vitesse » lorsque $v = +M$ alterne avec la « marche en arrière à toute vitesse » quand $v = -M$, etc. Il s'agit de choisir les signes de σ et la durée des intervalles de façon qu'on rejoigne à partir de l'état initial donné un certain état final imposé, par exemple, l'origine des coordonnées (fig. 3.2). Le problème de la sélection de tous ces paramètres sera examiné dans ce qui suit.

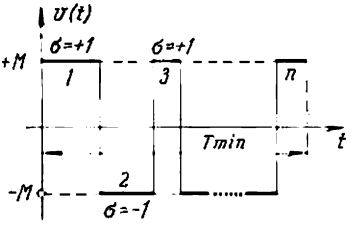


Fig. 3.4.

Démontrons le théorème des n intervalles dans le cas particulier quand la contrainte (3.14) est de la forme

$$\left|\frac{d^n x}{dt^n}\right| \leqslant M, \tag{3.31}$$

et l'équation (3.18) devient égalité

$$\frac{d^n x}{dt^n} = v. \tag{3.32}$$

La relation entre x et v est celle qu'on aurait pour une chaîne composée de n circuits intégrateurs (fig. 3.5), v étant sa grandeur

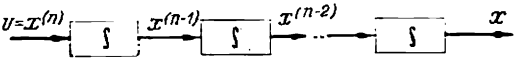


Fig. 3.5.

d'entrée et x sa grandeur de sortie. Appelons le schéma de la figure 3.5 schéma équivalent. Il coïncide avec le schéma réel de l'objet quand l'équation de l'objet est (3.32), c'est-à-dire

$$\frac{d^n x}{dt^n} = u. \tag{3.33}$$

Il est alors évident que $v = u$. Toutefois, dans le cas général, $v \neq u$. Montrons que pour obtenir le processus optimal $x(t)$, il faut fournir à l'entrée du schéma équivalent la grandeur $v = \sigma M$, où $\sigma = \pm 1$, les signes de σ alternent dans les intervalles adjacents et le nombre total des intervalles est n. Supposons qu'un tel processus existe et que sa durée est T_0. L'instant $t = T_0$ donne lieu aux égalités

$$x = x^*, \quad x^{(1)} = x^{*(1)}, \quad \ldots, \quad x^{(n-1)} = x^{*(n-1)}, \tag{3.34}$$

où $x^{(k)}$ et $x^{*(k)}$ sont les dérivées d'ordre k. Considérons d'abord les courbes $x^{(n-1)}$ et $x^{*(n-1)}$ (fig. 3.6).

D'après ce qui précède ces courbes se confondent pour $t \geqslant T_0$. La courbe $x^{(n-1)}(t)$ a la forme d'une droite brisée du fait qu'elle matérialise la grandeur de sortie du circuit intégrateur du schéma équivalent (fig. 3.5) à l'entrée duquel est appliquée la grandeur $v = \pm M$ variant par sauts (cf. fig. 3.4). La droite brisée $x^{(n-1)}$ comporte des segments à pente positive $+M$ qui alternent avec des segments à pente négative $-M$ de même module.

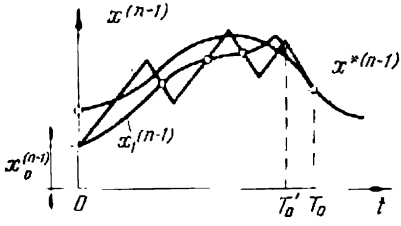


Fig. 3.6.

Supposons qu'il existe une autre courbe $x_1(t)$ ayant les mêmes conditions initiales que $x(t)$ mais donnant la durée du processus transitoire T_0' moindre, donc $T_0' < T_0$. Nous allons montrer que l'existence d'une courbe $x_1(t)$ ayant les propriétés indiquées et vérifiant la contrainte (3.31) est impossible. IO

Adjoignons à la courbe $x_1(t)$ un segment de la courbe $x^*(t)$ dans l'intervalle $T_0' \leqslant t \leqslant T_0$. Dans ce qui suit nous allons entendre par courbe $x_1(t)$ la courbe « prolongée », donnée dans l'intervalle $0 \leqslant t \leqslant T_0$. On en tire

$$x_1^{(k)}(T_0) = x^{(k)}(T_0) \quad (k = 0, 1, \ldots, n-1). \tag{3.35}$$

La courbe $x_1^{(n-1)}$ ne peut pas couper plus d'une fois chacun des segments de la ligne brisée $x^{(n-1)}$. En effet, s'il n'en était pas ainsi, sa dérivée $x_1^{(n)}$ serait d'un module supérieur par rapport à $x^{(n)} = \pm M$; mais dans ces conditions la courbe $x_1(t)$ serait inadmissible. Ensuite, la courbe $x_1^{(n-1)}$ ne peut pas couper ni le premier ni le dernier des segments (dans l'intervalle $0 \leqslant t \leqslant T_0$) de la ligne brisée $x^{(n-1)}$, car aux points extrêmes de ces segments les deux courbes se confondent ; dans le cas contraire la courbe $x_1(t)$ serait aussi inadmissible. Par conséquent, *le nombre de points d'intersection des courbes $x^{(n-1)}$ et $x_1^{(n-1)}$ ne peut dépasser $n - 2$.*

Considérons maintenant les courbes $x^{(n-2)}(t)$ et $x_1^{(n-2)}(t)$. Elles coïncident aux points $t = 0$ et $t = T_0$, tandis que leurs dérivées, comme nous l'avons montré ci-dessus, n'ont pas plus de $n - 2$ points d'intersection. On en déduit que les courbes $x^{(n-2)}$ et $x_1^{(n-2)}$ n'en ont pas plus de $n - 3$ à l'intérieur de l'intervalle $0 \leqslant t \leqslant T_0$ (on ne compte pas les points extrêmes de cet intervalle). En effet, si deux courbes continues et dérivables se coupent en deux points quelconques aux instants t_1 et t_2, $t_1 < t_2$, en un certain point intermédiaire leurs dérivées doivent être égales entre elles pour $t_1 < t < t_2$. Par conséquent, si les courbes $x^{(n-2)}$ et $x_1^{(n-2)}$ avaient plus de $n - 3$ points d'intersection, compte tenu de leur coïncidence aux extrémités de l'intervalle, on aurait plus de $n - 1$ points d'intersection de ces courbes. Il s'ensuit que leurs dérivées $x^{(n-1)}$ et $x_1^{(n-1)}$ se confondent en plus de $n - 2$ points à l'intérieur de l'intervalle $0 \leqslant t \leqslant T_0$; or, nous avons montré dans ce qui précède que ceci est impossible. Ainsi les courbes $x^{(n-2)}$ et $x_1^{(n-2)}$ ne se coupent pas plus qu'en $n - 3$ points à l'intérieur de l'intervalle $0 \leqslant t \leqslant T_0$.

En raisonnant d'une manière analogue on montre aisément que les courbes $x^{(n-3)}$ et $x_1^{(n-3)}$ n'ont pas plus de $n - 4$ points d'intersection à l'intérieur de l'intervalle $0 \leqslant t \leqslant T_0$, les courbes $x^{(n-4)}$ et $x_1^{(n-4)}$ pas plus de $n - 5$ points d'intersection, etc. En parcourant ainsi de gauche à droite les circuits du schéma équivalent de la figure 3.5 on aboutit enfin aux courbes $\frac{d^2x}{dt^2} = x^{(2)}$ et $\frac{d^2x_1}{dt^2} = x_1^{(2)}$. Il est clair que ces courbes n'ont pas plus d'un point d'intersection. On en déduit que les courbes $x^{(1)}$ et $x_1^{(1)}$ n'ont aucun point d'intersection à l'intérieur de l'intervalle $0 \leqslant t \leqslant T_0$. Mais ceci signifie que

$$x_1(T_0) = x_0 + \int_0^{T_0} x_1^{(1)}(dt) \neq x_0 + \int_0^{T_0} x^{(1)}(dt) = x(T_0), \qquad (3.36)$$

c est-à-dire que l'égalité (3.35) est compromise, bien qu elle devait avoir lieu si le temps T_0' de la courbe $x_1(t)$ était inférieur au temps T_0 de la courbe $x(t)$. Cette contradiction signifie qu'il ne peut exister aucune courbe admissible $x_1(t)$ avec $T_0' < T_0$. Par conséquent, la courbe $x(t)$ est celle d'un processus optimal et la démonstration du théorème des n intervalles pour la contrainte (3.31) est achevée.

Ces raisonnements sont aisément généralisés au cas de la contrainte (3.14) lorsque les circuits du schéma équivalent sont soit intégrateurs soit inertiels [3.5, 3.25] *).

Le théorème des n intervalles ne donne pas à lui seul la règle de sélection du signe de σ sur le premier intervalle ni de la durée des intervalles. Pourtant, ce théorème permet de réaliser la synthèse

*) C'est précisément le cas de la contrainte imposée aux racines de l'équation (3.17).

d'un schéma fonctionnel ou, en d'autres termes, la synthèse de l'algorithme de l'organe A de commande optimal qui résout automatiquement le problème [3.11]. Si l'on connaît la valeur de la fonction de commande v nécessaire à l'instant concerné, la valeur respective de la commande réelle u s'obtient sans peine. Si l'objet est soumis à la contrainte de type (3.16), on a $v = u$. Mais si l'équation de l'objet et l'équation (3.18) de la contrainte ne coïncident pas, on a $v \neq u$; pourtant la relation entre u et v s'obtient sans difficulté. Par exemple, dans le cas de l'équation (3.19) et de la contrainte (3.10), avec $(dx/dt)_{t=0} = 0$, on a

$$u^*(t) = b_0 v(t) + b_1 \int_0^1 v\,dt. \tag{3.37}$$

Le problème principal consiste donc à définir la fonction de commande v pour tout point de l'espace de phase $\bar{x}$ à l'instant t:

$$v = v(\bar{x}, t). \tag{3.38}$$

Pour trouver cette relation, considérons l'espace de phase des erreurs. En vertu du théorème des n intervalles la valeur de v en un point quelconque de l'espace de phase ne peut être autre que $+M$ ou $-M$, c'est-à-dire si $v = \sigma M$, on a $\sigma = +1$ ou $\sigma = -1$. Par conséquent, à chaque instant, à tout point ordinaire de l'espace de phase correspond soit $\sigma = +1$, soit $\sigma = -1$. Alors à chaque instant tout l'espace de phase est partagé en deux domaines caractérisés par les valeurs $\sigma = +1$ et $\sigma = -1$. La figure 3.7 représente à titre d'exemple un espace de phase tridimensionnel aux domaines $\sigma = +1$ et $\sigma = -1$. Ces domaines sont séparés par l'hypersurface S de dimension $n - 1$. Le but de la synthèse consiste précisément à décrire cette hypersurface en un instant quelconque. On appelle S *hypersurface de commutation*.

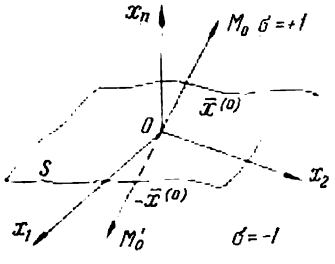


Fig. 3.7.

Dans le cas général l'hypersurface S a une forme différente à des instants différents. Les hypersurfaces de ce type sont dites *non stationnaires*. C'est le cas où le temps est présent explicitement dans les équations de l'objet. Dans une classe plus restreinte, l'hypersurface S est fixe, mais sa forme dépend des paramètres de la fonction de consigne $x^*(t)$. Une telle hypersurface se nomme *quasi stationnaire*.

Dans une classe encore plus restreinte, la forme de l'hypersurface S ne dépend pas du tout des paramètres de $x^*(t)$; on l'appelle alors

stationnaire. Dans ce dernier cas que nous allons étudier en détail l'équation de l'erreur n'est pas assujettie aux paramètres de $x^*(t)$.

Considérons deux points M_0 et M_0' dans un espace de phase des erreurs (fig. 3.7) pour le cas d'une hypersurface stationnaire S. Ces points sont symétriques par rapport à l'origine des coordonnées et correspondent aux valeurs $\bar{x}^{(0)}$ et $[-\bar{x}^{(0)}]$ du vecteur $\bar{x}$. Supposons que l'équation (3.18) est précisément celle des erreurs. Soit au point M_0 la valeur de σ égale à $+1$ et au point M_0' à -1. En effet, si le processus optimal $\bar{x}(t)$ vérifiant l'équation (3.18) démarre du point $\bar{x}^{(0)}$ et se déroule sous l'effet de l'action $v(t)$, le processus $-\bar{x}(t)$ soumis à l'action $-v(t)$ et partant du point $-\bar{x}^{(0)}$ vérifie également l'équation (3.18) et est aussi optimal.

Etant donné que n'importe quels points symétriques par rapport à l'origine des coordonnées, ne reposant pas sur l'hypersurface S, appartiennent à des domaines différents, l'hypersurface S passe par l'origine des coordonnées. Ensuite, on tire de la définition de l'hypersurface qu'elle ne comporte pas de « trous » par lesquels il serait possible de passer d'un domaine à l'autre sans couper l'hypersurface. Enfin, l'hypersurface S s'étend jusqu'aux points éloignés à l'infini de l'espace de phase.

Dans le cas général, dans un espace de phase, une trajectoire optimale est constituée de n portions qui correspondent aux intervalles $\sigma = \text{const}$. Par exemple, la figure 3.8 représente pour $n = 3$ la trajectoire de phase optimale M_0PQO qui part du point M_0 (vecteur $\bar{x}^{(0)}$) et compte trois portions. Pour la première portion M_0P, $\sigma = +1$; pour la deuxième PQ, $\sigma = -1$; enfin pour la troisième QO, $\sigma = +1$. La trajectoire $M_0'P'Q'O$ antisymétrique, ne figurant pas sur le dessin, part du point M_0' (vecteur $-\bar{x}^{(0)}$). Les valeurs de σ alternent dans l'ordre suivant: $\sigma = -1$; $+1$; -1. La dernière portion $Q'O$ de cette trajectoire, représentée en trait interrompu sur la figure 3.8, appartient au domaine où $\sigma = -1$ et mène le point représentatif à l'origine des coordonnées.

Soit le processus transitoire démarre à $t = 0$. On peut compter le temps « à rebours » en introduisant la variable $\tau = T - t$, où T est l'instant final du processus transitoire. Alors c'est la valeur $\tau = 0$ qui correspond à l'instant final du processus transitoire, c'est-à-dire à celui où le point représentatif est recalé à l'origine des coordonnées. A mesure que τ croît à partir du zéro, le point représentatif effectue un mouvement rétrograde suivant la trajectoire optimale, par exemple, suivant la trajectoire $OQPM_0$, en se déplaçant du point O vers Q, puis de Q vers P et de P vers M_0. Le point M_0 sera atteint à l'instant $\tau = T$ quand $t = 0$.

Il importe de noter que le point O ne peut être rejoint qu'en suivant l'une des deux trajectoires optimales possibles: QO ou $Q'O'$. En effet, introduisons $\tau = T - t$ et suivons l'une des trajectoires

« à rebours ». Sur la première portion à compter du point O seules deux valeurs de la fonction de commande v sont possibles : $+M$ ou $-M$, ce qui correspond aux valeurs $\sigma = +1$ et $\sigma = -1$. Avec $\sigma = +1$ le déplacement s'effectue suivant la trajectoire OQ et avec $\sigma = -1$ suivant OQ'. Désignons ces trajectoires par L_1' et L_1'' respectivement (fig. 3.9).

On peut rejoindre ces trajectoires de différentes façons. Un point *défini* quelconque de la trajectoire L_1' peut être atteint en suivant une trajectoire du type PQ associée à la valeur $\sigma = -1$. Désignons par

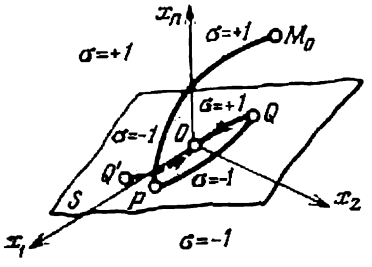


Fig. 3.8.

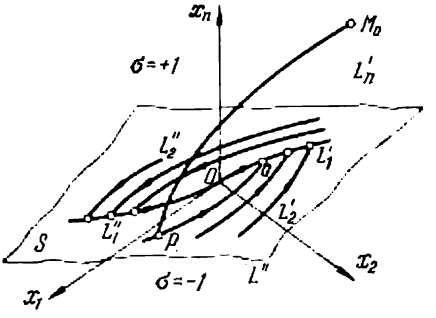


Fig. 3.9.

L_2' l'ensemble des points des trajectoires répondant à $\sigma = -1$ et menant aux points de la trajectoire L_1'. D'une façon analogue, à chaque point défini de la trajectoire appartenant à l'ensemble L_2' mène une seule trajectoire associée à la valeur $\sigma = +1$ (par exemple, PM_0 sur la figure 3.9). Désignons par L_3' l'ensemble des points des trajectoires associées à $\sigma = +1$ et conduisant à L_2', etc. Construisons d'une façon analogue les ensembles L_4', L_5', . . ., L_{n-1}' et enfin L_n'. Ce dernier ensemble inclut tous les points réguliers de l'espace de phase appartenant au domaine $\sigma = +1$ et ne se trouvant pas sur l'hypersurface S.

D'une manière analogue désignons par L_2'' l'ensemble des points des trajectoires qui mènent à L_1'' ; puis par L_3'' celui des points des trajectoires qui vont à L_2'', etc. Il est clair que L_n'' inclut tous les points réguliers de l'espace de phase appartenant au domaine $\sigma = -1$ et ne reposant pas sur l'hypersurface S.

La figure 3.9 montre qu'après le premier changement de signe de σ, les « jets » de dimension n des trajectoires de phase optimales deviennent des « jets » de dimension $n - 1$, puis après un second changement de signe, ils deviennent de dimension $n - 2$, etc. ; enfin au n-ième intervalle apparaît le « jet » unidimensionnel coulant vers l'origine des coordonnées. Il n'existe que deux jets unidimensionnels de ce type, L_1' et L_1''. Tout ce qui vient d'être dit s'observe aisé-

ment si l'on effectue le déplacement « à rebours » du point représentatif en partant de l'origine.

Le premier changement de signe a lieu au point P de la trajectoire optimale (si le temps t est considéré au sens direct). Le changement de signe de σ ne pouvant avoir lieu que lors de l'intersection de l'hypersurface de commutation S, le point P appartient à S. Etant donné que l'ensemble des points P est un ensemble des sous-ensembles L'_{n-1} et L''_{n-1} que nous allons noter L_{n-1}, ce dernier appartient à S. Mais tout point R de S peut être rejoint en suivant la trajectoire $\sigma = +1$ ou $\sigma = -1$ à partir d'un point régulier quelconque sans commutation, car la commutation ne se produit que sur l'hypersurface S. Par conséquent tout point R de l'hypersurface S appartient à l'ensemble L_{n-1}.

Ainsi, nous avons établi que l'hypersurface S de dimension $n - 1$ se confond avec l'ensemble L_{n-1} de dimension $(n - 1)$; par conséquent la recherche de S se ramène à celle des points de L_{n-1}.

Puisque dans le cas de l'hypersurface S stationnaire la famille des trajectoires L''_{n-1} est antisymétrique à L'_{n-1} et s'obtient en inversant les signes de toutes les coordonnées des points de L'_{n-1}, il suffit d'étudier seulement la construction de la famille L'_{n-1}. Pour réaliser cette opération, il est commode de se déplacer « à rebours » à partir de l'origine en remplaçant t par $\tau = T - t$. Le point représentatif se déplace d'abord suivant la trajectoire L'_1 (fig. 3.9) associée à la valeur $\sigma = +1$. Supposons qu'à l'instant τ_1 se produise le changement de signe de σ et le passage sur la trajectoire L'_2 ; à l'instant $\tau_2 > \tau_1$ on a un nouveau changement de signe de σ et le passage sur la trajectoire L'_3, etc. Enfin pour $\tau_{n-1} > \tau_{n-2} > \ldots > \tau_1$ on enregistre le « dernier » (le premier dans le temps réel) changement de signe de σ. En passant dans l'équation

$$a_0 \frac{d^n x}{dt^n} + a_1 \frac{d^{n-1}x}{dt^{n-1}} + \ldots + a_{n-1} \frac{dx}{dt} + a_n x = v \qquad (3.39)$$

au nouvel argument $\tau = T - t$, on obtient une nouvelle équation

$$a_0 (-1)^n \frac{d^n x}{d\tau^n} + a_1 (-1)^{n-1} \frac{d^{n-1}x}{d\tau^{n-1}} + \ldots$$

$$\ldots + (-1)\, a_{n-1} \frac{dx}{d\tau} + a_n x = v. \qquad (3.40)$$

Cherchons la solution de cette équation pour les conditions « initiales »

$$(x)_{\tau=0} = \left(\frac{dx}{d\tau}\right)_{\tau=0} = \ldots = \left(\frac{d^{n-1}x}{d\tau^{n-1}}\right)_{\tau=0} = 0 \qquad (3.41)$$

et avec $v = +M$, c'est-à-dire avec $\sigma = +1$. On obtient alors l'équation de la trajectoire L'_1 paramétrisée : $\bar{x} = \bar{x}(\tau)$. Pour $\tau = \tau_1$ le point de l'espace de phase assujetti à cette trajectoire a les coordonnées $\bar{x}(\tau_1)$. Si à cet instant on remplace $v = +M$ par $v = -M$,

la solution de l'équation (3.40) à l'instant $\tau_2 > \tau_1$ est une fonction de τ_1 de même que de τ_2 : $\bar{x} = \bar{x}(\tau_1, \tau_2)$.

En effectuant les changements de signe de σ aux instants $\tau_1, \tau_2, \ldots, \tau_{n-1}$ on obtient la solution de l'équation (3.40) sous la forme

$$\bar{x} = \bar{x}(\tau_1, \tau_2, \ldots, \tau_{n-1}). \tag{3.42}$$

Les équations des coordonnées $x_1, x_2, \ldots, x_n$ du vecteur $\bar{x}$ s'écrivent

$$x_i = x_i(\tau_1, \tau_2, \ldots, \tau_{n-1}) \quad (i = 1, 2, \ldots, n). \tag{3.43}$$

Ces équations donnent dans leur ensemble l'équation paramétrique de l'hypersurface S et par là même la solution du problème de la synthèse. En éliminant de l'équation (3.43) les paramètres $\tau_1, \tau_2, \ldots, \tau_{n-1}$ on peut obtenir dans plusieurs cas l'équation de S sous une forme explicite, c'est-à-dire sous la forme d'une relation associant les coordonnées $x_1, x_2, \ldots, x_n$:

$$\psi(x_1, x_2, \ldots, x_n) = 0. \tag{3.44}$$

Soit la fonction ψ positive d'un côté de l'hypersurface S et négative de l'autre côté. Soit, par exemple, $\psi > 0$ pour les points du domaine $\sigma = +1$ et $\psi < 0$ pour ceux du domaine $\sigma = -1$. On peut poser alors

$$\sigma = \operatorname{sign} \psi \tag{3.45}$$

et

$$v = \sigma M = M \operatorname{sign} \psi = M \operatorname{sign} \psi(x_1, x_2, \ldots, x_n). \tag{3.46}$$

C'est précisément l'algorithme recherché du fonctionnement d'un système optimal, c'est-à-dire l'équation $v = v(\bar{x})$. A tout point $\bar{x}$ de l'espace de phase l'équation (3.46) associe une valeur de v qui correspond à la loi de commande optimale. Des résultats analogues peuvent être obtenus pour le cas d'une surface S quasi stationnaire [3.11, 3.25], à cette différence près que dans ce cas L'_{n-1} et L''_{n-1} ne sont pas antisymétriques.

Connaissant la relation (3.46) on peut construire un schéma fonctionnel de l'organe de commande optimal (fig. 3.10). La grandeur de sortie X de l'objet B est amenée par la chaîne de réaction OO' à l'entrée du bloc Σ de l'organe de commande A. La grandeur X peut être un vecteur à plusieurs coordonnées. La chaîne de réaction est composée alors de plusieurs circuits qui transmettent les coordonnées X_i du vecteur $\bar{X}$, par exemple X, $\frac{dX}{dt}$, etc.

On amène également au bloc Σ le vecteur $\bar{X}^*$ de la consigne. Le bloc Σ élabore les coordonnées $x_1, x_2, \ldots, x_n$ du vecteur de l'erreur : $x_i = X_i^* - X_i$. Au besoin on adjoint au bloc Σ des dérivateurs pour définir les coordonnées faisant défaut par dérivation de certaines coordonnées d'entrée. Les erreurs $x_1, x_2, \ldots, x_n$ sont appliquées

sur l'entrée du convertisseur non linéaire CNL qui produit leur fonction non linéaire $\psi(x_1, x_2, \ldots, x_n)$. Cette dernière est dirigée sur l'entrée du circuit à relais CR dont la grandeur de sortie $M \operatorname{sign} \psi$ est justement la grandeur v. Cette grandeur est amenée à l'entrée de l'organe A' qui transforme $v(t)$ en commande $u(t)$. Si $u(t) = v(t)$, l'organe A' est inutile.

Il faut expliquer comment en présence d'une seule hypersurface de commutation S de dimension $n - 1$ on réalise dans un espace de

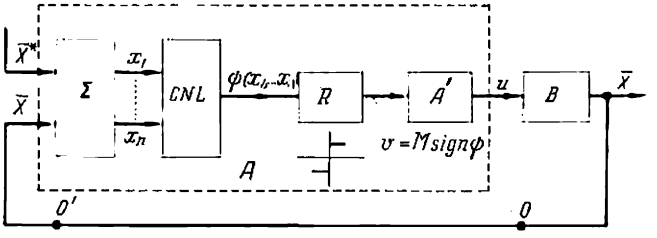


Fig. 3.10.

phase de dimension n un processus composé de n intervalles. Pour le comprendre il faut retenir que n'importe quel système réel donne lieu à des fluctuations qui « repoussent » légèrement le point représentatif de la trajectoire de phase optimale. C'est pourquoi le mouvement réel d'un système dépend non seulement de la forme de

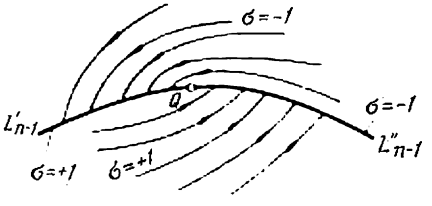


Fig. 3.11.

l'hypersurface S mais encore de la structure de l'espace de phase dans le voisinage de cette hypersurface. Cette structure est schématisée sur la figure 3.11. La partie L'_{n-1} de l'hypersurface S correspond ici à la valeur $\sigma = +1$, et L''_{n-1} à la valeur $\sigma = -1$. D'un côté de L'_{n-1} (en bas de la fig. 3.11) passent presque parallèlement d'autres trajectoires qui appartiennent au domaine $\sigma = +1$ de l'espace de phase. De l'autre côté de L'_{n-1} (en haut de la figure 3.11) passent les trajectoires $\sigma = -1$ qui conduisent précisément à cet ensemble L'_{n-1}. Comme on le voit sur la figure 3.11, dans le voisinage

de la partie L''_{n-1} de l'hypersurface S la situation est analogue à cette différence près que ce sont les trajectoires $\sigma = +1$ qui mènent à L''_{n-1}.

Si le point représentatif qui suit la trajectoire L'_{n-1} est chassé dans une zone voisine du domaine $\sigma = -1$ de l'espace de phase, σ change de signe. Mais ensuite le point représentatif revient immédiatement (dans le cas d'une fluctuation suffisamment petite) à l'hypersurface. Dans ce cas le mouvement réel du point ne change pas.

Il en est tout autrement (et une telle situation est inévitable du fait de l'existence dans le système des fluctuations les plus différentes) si le point représentatif est chassé dans le domaine $\sigma = +1$ voisin de la partie L'_{n-1} de l'hypersurface S. Le point concerné ne revient pas alors immédiatement sur l'hypersurface. Sa trajectoire ultérieure passe par le domaine $\sigma = +1$ dans le voisinage immédiat de L'_{n-1}. Par conséquent la trajectoire sera *quasi* la même que si le point représentatif suivait l'hypersurface. Au bout d'un intervalle de temps fini le point représentatif arrive à la partie L''_{n-1} de l'hypersurface S dans le voisinage immédiat du point Q auquel aboutit la trajectoire « idéale » de L'_{n-1}. Si maintenant une nouvelle fluctuation repousse le point représentatif, il s'engage sur la trajectoire $\sigma = -1$ proche de l'une des trajectoires de L'_{n-2}, car c'est par Q que passe la trajectoire appartenant à cet ensemble. Après n mouvements de ce type, le point représentatif, en se déplaçant tour à tour dans les domaines $\sigma = -1$, $\sigma = +1$, tombe dans le voisinage immédiat de l'origine des coordonnées et de sorte que le processus transitoire prend fin.

Par conséquent, la trajectoire réelle passe *près* de l'hypersurface S et à chaque commutation le point représentatif « perce » l'hypersurface S. La figure 3.12 montre une trajectoire réelle dans un espace tridimensionnel ($n = 3$). La trajectoire idéale est représentée par la courbe M_0PQO dont les portions PQ et QO reposent sur la surface de commutation S. Si dans le voisinage de la position P une fluctuation chasse le point représentatif de l'hypersurface S dans le domaine $\sigma = -1$, le point vient en position P_1 et se déplace ensuite suivant la trajectoire $P_1Q'_1$ voisine de la trajectoire idéale PQ. En Q'_1 le point représentatif atteint la surface S. Si maintenant une nouvelle fluctuation chasse le point en position Q_1 avec $\sigma = +1$, il suivra la trajectoire Q_1O_1 voisine de la trajectoire « idéale » QO. En O_1 le point représentatif tombe dans le voisinage immédiat de l'origine des coordonnées et on peut considérer que le processus transitoire est achevé.

Revenons encore une fois à la construction de l'hypersurface S et rappelons qu'elle représente le lieu géométrique des points de *premier* changement de signe de σ. Pourtant tous les $n - 1$ autres changements de signe de σ se produisent sur cette même hypersurface. En effet, il résulte de ce qui vient d'être dit que pour assurer les changements ultérieurs de signe de σ aucun besoin n'est de créer des sur-

faces ou des lignes quelconques à l'intérieur de L_{n-1}. Il suffit d'avoir à sa disposition une seule « cloison » dans un espace de phase de dimension n correspondant au premier changement de signe. Grâce aux fluctuations le processus réel peut être aussi proche que l'on veut du processus optimal idéal mais jamais identique. Dans un système

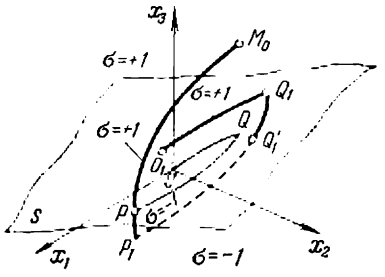


Fig. 3.12.

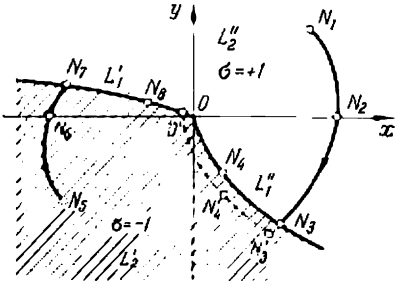


Fig. 3.13.

réel le processus idéal est impossible car la probabilité pour le point représentatif de se trouver sur l'hypersurface S et de se déplacer sur cette surface est égale à zéro.

Montrons sur un exemple très simple d'un objet de deuxième ordre comment on définit la fonction ψ et l'hypersurface S. Dans le cas concerné l'espace de phase des erreurs devient un plan de phase (fig. 3.13) et la surface dégénère en une courbe de commutation qui coupe l'espace de phase en deux domaines: $\sigma = +1$ et $\sigma = -1$. Soit l'objet B décrit par l'équation

$$a_0 \frac{d^2x_1}{dt^2} = u, \tag{3.47}$$

et la contrainte de la forme

$$|u| \leqslant M. \tag{3.48}$$

La classe des consignes est de la forme

$$x_1^*(t) = A_0 + A_1 t + A_2 t^2, \tag{3.49}$$

où A_0, A_1 et A_2 sont des grandeurs constantes pour chaque processus particulier. Les fonctions admissibles $x_1^*(t)$ sont les solutions de l'équation (3.47) sous la condition

$$|u| < M. \tag{3.50}$$

Le signe $\leqslant$ est remplacé ici par $<$ pour que le point $\bar{x}$ puisse « rattraper » le point $\bar{x}^*$. Etant donné que

$$a_0 \left| \frac{d^2x_1^*}{dt^2} \right| < M, \tag{3.51}$$

en portant la seconde dérivée de l'expression (3.49) dans la condition (3.51), on a

$$|A_2| < \frac{M}{2a_0}. \tag{3.52}$$

Cette condition est une contrainte sur les fonctions admissibles du type (3.49).

L'équation du processus optimal devient alors

$$\frac{d^2x_1}{dt^2} = \frac{\sigma M}{a_0}, \tag{3.53}$$

où $\sigma = \pm 1$.

Désignons par x l'erreur, c'est-à-dire la différence

$$x = x_1^* - x_1, \tag{3.54}$$

et par y sa dérivée :

$$y = \frac{dx}{dt} = \frac{dx_1^*}{dt} - \frac{dx_1}{dt}. \tag{3.55}$$

Les équations (3.53), (3.54) et (3.49) entraînent

$$\frac{d^2x}{dt^2} = \frac{d^2x_1^*}{dt^2} - \frac{d^2x_1}{dt^2} = 2A_2 - \frac{\sigma M}{a_0} = \mu_0. \tag{3.56}$$

Dans chaque intervalle la grandeur μ_0 est constante du fait que $\sigma = \text{const}$. Passons maintenant au temps « rétrograde » $\tau = T - t$. Compte tenu de l'équation (3.56), il vient

$$\frac{d^2x}{dt^2} = \frac{d}{dt}\left(\frac{dx}{dt}\right) = -\frac{d}{d\tau}\left(-\frac{dx}{d\tau}\right) = \frac{d^2x}{d\tau^2} = \mu_0. \tag{3.57}$$

Ensuite

$$y = \frac{dx}{dt} = -\frac{dx}{d\tau} = -\int\left(\frac{d^2x}{d\tau^2}\right)d\tau = -\mu_1 - \mu_0\tau \tag{3.58}$$

et

$$x = -\int y\,d\tau = \mu_2 + \mu_1\tau + \mu_0\frac{\tau^2}{2}, \tag{3.59}$$

où μ_1 et μ_2 sont des constantes. Recherchons les courbes L_1' et L_2'' qui constituent dans le cas considéré la courbe de commutation. Étant donné que pour $\tau = 0$ les grandeurs y et x associées à ces courbes deviennent nulles, (3.58) et (3.59) conduisent à $\mu_1 = \mu_2 = 0$. Par conséquent

$$x = \mu_0\frac{\tau^2}{2}, \quad y = -\mu_0\tau. \tag{3.60}$$

Cherchons d'abord la courbe L_1'' du domaine $\sigma = -1$. En vertu de (3.56) et de la condition (3.52) le signe de μ_0 est défini par celui de σ :

$$\operatorname{sign}\mu_0 = -\operatorname{sign}\sigma. \tag{3.61}$$

Etant donné que $\tau > 0$, on tire de la deuxième équation (3.60) que

$$\operatorname{sign} y = -\operatorname{sign} \mu_0. \tag{3.62}$$

Par conséquent

$$\operatorname{sign} y = \operatorname{sign} \sigma = \sigma. \tag{3.63}$$

Ainsi avec $\sigma = -1$ on a $\mu_0 > 0$ et $y < 0$. De ce fait le déplacement suivant la courbe L_1'' a lieu dans le demi-plan inférieur (fig. 3.13) En éliminant des deux équations (3.60) l'argument τ, on obtient l'équation de la courbe L_1'' sous la forme

$$\left.\begin{aligned} x &= \frac{y^2}{2\mu_0} > 0, \\ \sigma &= \operatorname{sign} \sigma = \operatorname{sign} y = -1. \end{aligned}\right\} \tag{3.64}$$

Pour la courbe L_1' on a $\sigma = +1$. Par conséquent, comme on le voit de (3.56), sign $\mu_0 = -1$ et en vertu de (3.62) $y > 0$. Dans ce cas les équations (3.60) donnent

$$\left.\begin{aligned} x &= \frac{y^2}{2\mu_0} < 0, \\ \sigma &= \operatorname{sign} \sigma = \operatorname{sign} y = +1. \end{aligned}\right\} \tag{3.65}$$

On peut réunir les équations des deux courbes L_1' et L_1'' en une seule pour la courbe $L_1 = S$:

$$x = \frac{y^2}{2\mu_0} = \frac{y^2}{2\left(2A_2 - \frac{M_0}{a_0} \operatorname{sign} y\right)}, \tag{3.66}$$

où au lieu de $\sigma = \operatorname{sign} \sigma$ figure sign y égale à σ. L'expression (3.66) est l'équation de la courbe de commutation. On peut l'écrire sous une forme analogue à (3.44), c'est-à-dire sous la forme $\psi = 0$, où

$$\psi(x, y) = x + \frac{y^2}{2\left(\frac{M}{a_0} \operatorname{sign} y - 2A_2\right)} = x + \frac{y^2 \operatorname{sign} y}{2\left(\frac{M}{a_0} - 2A_2 \operatorname{sign} y\right)}. \tag{3.67}$$

La figure 3.13 montre la courbe de commutation $N_3N_4ON_8N_7$ et les trajectoires optimales dans le plan de phase. Le point représentatif qui part de N_1 suit la trajectoire parabolique $N_1N_2N_3$ qu'on obtient aisément en cherchant la solution de l'équation (3.56) pour $\sigma = +1$ et les conditions initiales associées au point N_1. En N_3 le point représentatif atteint la courbe de commutation. Après le changement de signe, dans le cas idéal, le point représentatif se déplace le long de la courbe de commutation N_3N_4O vers l'origine des coordonnées. La présence des fluctuations fait qu'en réalité le point représentatif se déplace suivant le trait interrompu $N_3'N_4'O$ dans le domaine $\sigma = -1$, c'est-à-dire suivant une trajectoire voisine de N_3N_4O pour venir en O' dans le voisinage immédiat de l'origine.

Si à l'instant initial le point représentatif est en N_5, c'est-à-dire dans le domaine hachuré où $\sigma = -1$, il suit alors la trajectoire parabolique $N_5N_6N_7$ (solution de l'équation (3.56) pour $\sigma = -1$) jusqu'à la portion de la courbe de commutation N_7N_8O qui se situe dans le deuxième quadrant, puis la courbe de commutation vers l'origine des coordonnées.

Comme $\psi(x, y)$ (3.67) dépend du paramètre A_2 de la consigne $x_1^*(t)$, d'après la définition donnée dans ce qui précède la courbe de commutation S est *quasi stationnaire*. On pouvait s'y attendre car l'équation (3.56) de l'erreur $x(t)$ comporte le paramètre A_2. Dans ce cas les parties L_1' et L_1'' de la courbe de commutation S ne sont plus antisymétriques par rapport à l'origine. En effet, dans la formule (3.66) avec $y < 0$ le module du dénominateur est $\left|\frac{M}{a_0} + 2A_2\right|$, alors qu'avec $y > 0$ le module devient $\left|\frac{M}{a_0} - 2A_2\right|$. Par conséquent les demi-paraboles L_1' et L_1'' diffèrent et ne peuvent pas être amenées en coïncidence par rotation de l'une d'elles de 180°. Si l'on examine une classe des consignes plus restreinte

$$x_1^*(t) = A_0 + A_1 t, \tag{3.68}$$

pour laquelle $A_2 = 0$, les paramètres de la consigne ne figurent pas dans l'équation (3.56) de l'erreur $x(\tau)$. Dans ce cas l'expression (3.67) s'écrit

$$\psi(x, y) = x + \frac{y^2 \operatorname{sign} y}{\frac{2M}{a_0}} \tag{3.69}$$

et les deux demi-paraboles L_1' et L_1'' sont antisymétriques par rapport à l'origine. Le deuxième terme du second membre de la formule (3.69) représente la réaction quadratique en vitesse qui a déjà figuré dans [3.1].

Connaissant l'équation (3.67) on peut réaliser la synthèse d'un système optimal. En substituant à x dans l'expression de ψ une valeur suffisamment grande on obtient $\psi > 0$. Par conséquent la fonction $\psi > 0$ est au-dessus de la courbe de commutation. D'une façon analogue on montre sans peine que $\psi < 0$ est au-dessous de cette courbe. On peut appliquer alors la grandeur $w = \psi$ à l'entrée d'un circuit à relais dont la caractéristique est montrée à la figure 3.14,*a*. La grandeur de sortie u de ce circuit, égale à $\pm M$, peut être appliquée à l'entrée de l'objet. Cependant en pratique on est souvent obligé de remplacer ce circuit à relais par un circuit plus compliqué. En effet, la solution trouvée ci-dessus et l'organe de commande optimal correspondant ne présentent un avantage que dans le cas de grandes valeurs de l'erreur et de ses dérivées. Lorsque dans l'espace de phase des erreurs le point représentatif est dans le voisinage de l'origine

ou même de l'hypersurface S, la loi optimale de déplacement peut être remplacée par une autre, par exemple par une loi linéaire. La durée totale du processus transitoire augmente peu par suite d'un tel remplacement. Par contre on considère que les oscillations dans le voisinage de l'origine sont généralement indésirables. On peut

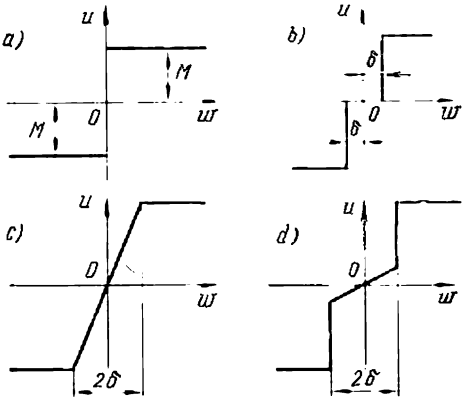


Fig. 3.14.

les diminuer en donnant à la relation $u = f(w)$ la forme d'une des courbes représentées sur les figures 3.14, *b*, *c*, *d*. Avec $|w| > \delta$, où δ est petit, ces relations se confondent avec la caractéristique d'un circuit à relais.

Le signal de commande w peut être remplacé par une autre grandeur quelconque w_1 de même signe, par exemple

$$w_1 = 2w\left(\frac{M}{a_0} - 2A_2 \operatorname{sign} y\right) =$$
$$= 2x\left(\frac{M}{a_0} - 2A_2 \operatorname{sign} y\right) + y^2 \operatorname{sign} y. \qquad (3.70)$$

Cette expression est plus commode que (3.67) car elle est exempte d'opération de division.

La figure 3.15 représente le schéma fonctionnel d'un système qui réalise la loi de commande (3.70). L'objet B est figuré ici par un double trait. La grandeur de sortie X de l'objet, qui est la grandeur à régler, après le changement de signe effectué par l'inverseur (-1), est amenée à l'entrée de l'additionneur $\sum$ qui reçoit également la consigne X_0. La différence $X_0 - X = x$ est fournie à l'entrée du dérivateur D_1 de même qu'à l'une des entrées du multiplieur M-2. La grandeur $M/a_0 - 2A_2$ sign y est amenée à la deuxième entrée de ce même multiplieur qui délivre à sa sortie le produit

$2x\,(M/a_0 - 2A_2 \operatorname{sign} y)$, c'est-à-dire le premier terme de la formule (3.70). Le facteur $M/a_0 - 2A_2 \operatorname{sign} y$ de cette expression s'obtient par l'addition de la grandeur $M/a_0 = \text{const}$ et de la sortie d'un autre multiplieur M-1. A l'entrée de ce dernier sont appliqués également les facteurs sign y et $-2A_2$. Le premier de ces facteurs est fourni par la sortie du convertisseur non linéaire (CNL-2) à caractéristique de relais recevant à son entrée la grandeur y. Le facteur $-2A_2$ est

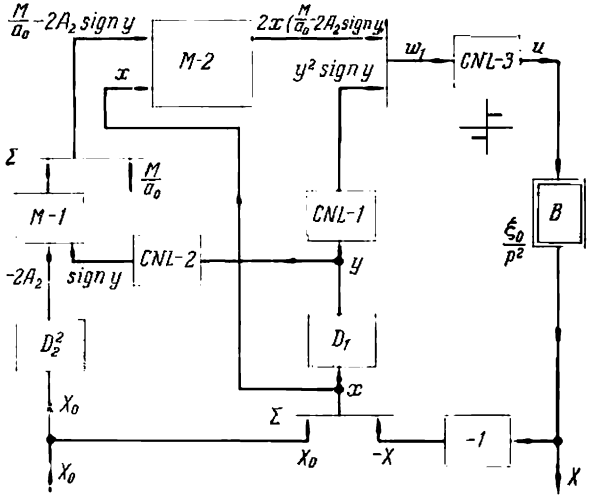


Fig. 3.15.

l'accélération $-d^2X_0/dt^2$ débitée en permanence par le double dérivateur D_2^2. On suppose que la consigne $X_0\,(t)$ n'est pas déformée par les bruits et la dérivation se produit sans distorsions.

La grandeur d'entrée du convertisseur non linéaire (CNL-3) est la somme figurant au second membre de l'égalité (3.70). Le deuxième terme de cette somme est fourni par le convertisseur non linéaire (CNL-1) qui reçoit y à son entrée. La caractéristique du convertisseur (CNL-3) est de l'un des types représentés sur la figure 3.14. En plus des organes représentés sur la figure 3.15, le schéma peut comporter des circuits supplémentaires nécessaires dans le cas de faibles écarts lorsque le système est régi par une loi autre que la loi de commande optimale [3.25].

L'algorithme (3.60) est relativement simple et dans le cas (3.69), lorsque $A_2 = 0$, il devient tout à fait trivial. Plus généralement, dans le cas d'une contrainte de deuxième ordre, sous la condition d'applicabilité du théorème des n intervalles, les algorithmes des systèmes strictement optimaux sont relativement simples. Soit,

par exemple, l'objet composé d'un circuit avec inertie et d'un intégrateur couplés en série ayant la fonction de transfert

$$K(p)=\frac{\xi_0}{p(1+pT)}=\frac{\bar{x}_1(p)}{\bar{u}(p)}, \tag{3.71}$$

où $\bar{x}_1(p)$ est la transformée de la grandeur à régler $x_1(t)$ à la sortie de l'objet et $\bar{u}(p)$ la transformée de $u(t)$.

Supposons que

$$|u(t)| \leqslant M_2 \tag{3.72}$$

et la consigne appartienne à la classe

$$x^*(t) = A_0 + A_1 t. \tag{3.73}$$

On montre aisément que les fonctions admissibles de cette classe doivent, dans le cas considéré, vérifier la condition

$$|A_1| < \xi_0 M_2 = M. \tag{3.74}$$

La formule de la grandeur w fournie à l'entrée du circuit à relais de l'équation

$$u(t) = M_2 \operatorname{sign} w \tag{3.75}$$

s'écrit alors

$$w=x+T\left[(A_1+M\operatorname{sign} y)\ln\left(1-\frac{y}{A_1+M\operatorname{sign} y}\right)+y\right], \tag{3.76}$$

où $x = x^* - x_1$ est l'erreur et y sa dérivée (cf. [3.25)].

Avec $A_1 = 0$ cette formule se met sous une forme très simple

$$w=x+T\left[M\operatorname{sign} y\ln\left(1-\frac{|y|}{M}\right)+y\right]. \tag{3.77}$$

La courbe de commutation donnée par la formule (3.76) dépend du paramètre A_1 de la consigne. Elle est donc quasi stationnaire. Dans le cas particulier qui correspond à l'expression (3.77) la courbe de commutation ne dépend pas des paramètres de la consigne et est stationnaire.

Considérons à titre d'exemple le problème des trajectoires dans un espace de phase tridimensionnel. On borne en valeur absolue la dérivée d'ordre trois de la grandeur x_1 à régler

$$\left|\frac{d^3x_1}{dt^3}\right| \leqslant M. \tag{3.78}$$

On peut montrer que, si la consigne x^* appartient à la classe des paraboles (3.49), la surface de commutation S est stationnaire. C'est pourquoi nous pouvons poser pour plus de commodité que $x^* = A_0 = \text{const}$, ce qui ne modifie pas le résultat, car de toutes façons S est indépendante des paramètres de la fonction $x^*(t)$ [3.11, 3.25].

L'équation de la contrainte (3.39) s'écrit

$$\frac{d^3x_1}{dt^3} = \sigma M = v. \tag{3.79}$$

Avec les notations du temps « rétrograde » $\tau = T - t$ elle prend la forme

$$\frac{d^3x_1}{d\tau^3} = -\sigma M. \tag{3.80}$$

Cherchons à résoudre cette équation pour les conditions initiales nulles (3.41) et $\sigma = \text{const}$. Posons

$$\left.\begin{aligned} x = x^* - x_1,\ y = \frac{dx}{dt} = -\frac{dx}{d\tau}, \\ z = \frac{d^2x}{dt^2} = \frac{d^2x}{d\tau^2} = -\frac{dy}{d\tau}. \end{aligned}\right\} \tag{3.81}$$

Alors (3.70) entraîne avec $\sigma = +1$

$$\frac{dz}{d\tau} = -M, \tag{3.82}$$

d'où pour $\tau = \tau_1$ il vient

$$z = -M\tau_1. \tag{3.83}$$

Ensuite on a pour $\tau = \tau_1$

$$y = -\int_0^{\tau_1} z\,d\tau = M\frac{\tau_1^2}{2} \tag{3.84}$$

et

$$x = -\int_0^{\tau_1} y\,d\tau = -M\frac{\tau_1^3}{6}. \tag{3.85}$$

Les égalités (3.83), (3.84) et (3.85) sont les équations paramétrisées de L_1'. Ces égalités donnent à l'instant $\tau = \tau_1$ les coordonnées

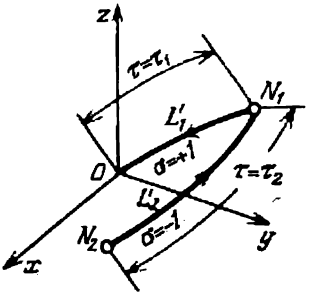


Fig. 3.16.

du point représentatif N_1 issu de l'origine des coordonnées (fig. 3.16). Soit maintenant à cet instant la grandeur σ qui change pour devenir

$\sigma = -1$. On a alors à intégrer l'équation

$$\frac{dz}{d\tau} = +M \tag{3.86}$$

sous les conditions initiales (3.83)-(3.85). Si la lecture du temps τ_2 débute à l'instant de commutation, les formules de z, y, x pour la deuxième portion se mettent sous la forme

$$z = \int_0^{\tau} M\, d\tau_2 - M\tau_1 = M(\tau_2 - \tau_1), \tag{3.87}$$

$$y = -\int_0^{\tau_2} z\, d\tau_2 + M\frac{\tau_1^2}{2} =$$

$$= -\int_0^{\tau_2} M(\tau_2 - \tau_1)\, d\tau_2 + \frac{M\tau_1^2}{2} = M\tau_1^2 - \frac{M}{2}(\tau_2 - \tau_1)^2 \tag{3.88}$$

et

$$x = -\int_0^{\tau_2} y\, d\tau_2 - M\frac{\tau_1^3}{6} = -\int_0^{\tau_2}\left[M\tau_1^2 - \frac{M}{2}(\tau_2 - \tau_1)^2\right] d\tau_2 - M\frac{\tau_1^3}{6} =$$

$$= -M\tau_1^2\tau_2 + \frac{M}{6}(\tau_2 - \tau_1)^3. \tag{3.89}$$

Les expressions (3.87), (3.88) et (3.89) sont les équations paramétrisées de la surface L_2', celles de la surface L_2'' s'obtiennent aisément en inversant le signe de x, y, z

$$\left.\begin{aligned} x &= M\tau_1^2\tau_2 - \frac{M}{6}(\tau_2 - \tau_1)^3, \\ y &= -M\tau_1^2 + \frac{M}{2}(\tau_2 - \tau_1)^2, \\ z &= -M(\tau_2 - \tau_1). \end{aligned}\right\} \tag{3.90}$$

En modifiant les valeurs des paramètres τ_1 et τ_2 des formules précédentes dans des limites

$$\left.\begin{aligned} 0 \leqslant \tau_1 < \infty, \\ 0 \leqslant \tau_2 < \infty, \end{aligned}\right\} \tag{3.91}$$

on peut obtenir tous les points de la surface de commutation S. Dans l'exemple considéré on peut éliminer également τ_1 et τ_2 des deux groupes d'équations et, en réunissant les deux « demi-surfaces » L_1' et L_2'' [3.11, 3.25], obtenir une équation de la surface S sous une

forme explicite :

$$\psi(x, y, z) = x + \frac{z^3}{3M^2} + (\operatorname{sign} w)\left[\frac{yz}{M} + \sqrt{\frac{1}{M}}\left(\frac{z^2}{2M} + y \operatorname{sign} w\right)^{3/2}\right] = 0,$$

$$w = y + \frac{z^2}{2M} \operatorname{sign} z. \tag{3.92}$$

La fonction de commande s'exprime par la formule

$$v = M \operatorname{sign} \psi. \tag{3.93}$$

L'équation (3.92) est déjà relativement complexe. Pour d'autres exemples de contraintes de troisième ordre, les algorithmes d'un système strictement optimal le sont encore plus. Avec $n = 2$, les algorithmes sont aussi relativement compliqués si l'équation caractéristique compte des racines complexes conjuguées [3.23]. Pourtant, pour n'importe quels types de systèmes de troisième ordre on peut trouver des algorithmes approchés optimaux relativement simples (cf., par exemple, [3.67]).

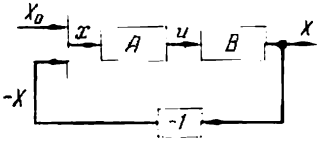
§ 2. Application des méthodes variationnelles classiques

Commençons par des exemples simples.

Soit l'objet défini par l'équation linéaire à coefficients constants

$$b_m \frac{d^m X}{dt^m} + \ldots + b_0 X = u(t), \tag{3.94}$$

où $X(t)$ est la grandeur de sortie et $u(t)$ la commande. La fonction de transfert de l'objet s'écrit



$$K_B(p) = \frac{\overline{X}(p)}{\overline{u}(p)} = \frac{1}{b_m p^m + \ldots + b_0} = \frac{1}{P_B(p)}. \tag{3.95}$$

Fig. 3.17.

La figure 3.17 représente le schéma fonctionnel d'un système asservi dans lequel l'erreur x constitue la différence entre la consigne X_0 et la grandeur à régler X :

$$x = X_0 - X.$$

Soit pour plus de simplicité

$$X_0 = \begin{cases} A_0 = \text{const}, & t \geqslant 0, \\ 0, & t < 0, \end{cases}$$

on considère ainsi la réponse transitoire du système à l'écart initial ; par ailleurs, pour $t < 0$ tous les circuits du système sont au repos et, par conséquent, la grandeur X, de même que ses dérivées, est égale

à zéro. L'erreur x est donnée par la somme

$$x = x_s + x_d, \tag{3.96}$$

où x_s est l'erreur du processus stationnaire (constante dans l'exemple considéré), alors que x_d est l'erreur transitoire qui s'annule une fois que le processus transitoire prend fin. On doit choisir un organe de commande A tel qu'il minimise l'intégrale

$$I_V = \int_0^\infty V\,dt = \int_0^\infty \left[x_d^2 + \gamma_1 \left(\frac{dx_d}{dt}\right)^2 + \ldots + \gamma_{n-1}\left(\frac{d^{n-1}x_d}{dt^{n-1}}\right)^2\right] dt. \tag{3.97}$$

Les problèmes relatifs aux objets linéaires où les critères d'optimalité sont donnés sous la forme d'une intégrale de la forme quadratique constituent le type le plus simple de problèmes de synthèse des systèmes optimaux, car dans ce cas l'organe de commande optimal est également linéaire. En effet, nous avons montré dans le chapitre II que la trajectoire optimale $x_d^*(t)$ qui minimise l'intégrale (3.97) est une solution de l'équation linéaire

$$\vartheta_n \frac{d^n x_d^*(t)}{dt^n} + \vartheta_{n-1} \frac{d^{n-1}x_d^*(t)}{dt^{n-1}} + \ldots + \vartheta_0 x_d^*(t) = 0. \tag{3.98}$$

Les coefficients ϑ_i de cette équation sont associés aux coefficients γ_i de la forme quadratique V de (3.97) par des relations qui s'obtiennent aisément à partir de l'équation d'Euler. Nous les avons données dans le chapitre II.

Dans le problème simple examiné dans ce qui suit les conditions initiales de $x_d(t)$ ne varient pas. Les cas des conditions initiales variables peuvent être étudiés par des méthodes classiques du calcul des variations. Il est pratiquement plus commode alors de considérer non pas l'équation générale du système mais l'ensemble des équations de ses circuits.

Etant donné que l'équation du système est linéaire d'ordre n et que l'équation de l'objet d'ordre $m < n$ est également linéaire, l'équation de l'organe de commande doit être aussi linéaire et son ordre est $n - m$. Soit $K_A(p)$ la fonction de transfert de cet organe. La fonction de transfert du système en boucle fermée s'écrit alors

$$K_f(p) = \frac{K_A(p)\,K_B(p)}{1 + K_A(p)\,K_B(p)} = \frac{\overline{X}(p)}{\overline{X}_0(p)}, \tag{3.99}$$

où $\overline{X}_0(p)$ est la transformée de la consigne. Compte tenu de (3.96) il vient

$$\overline{X}(p) = A_0 \frac{K_A(p)\,K_B(p)}{1 + K_A(p)\,K_B(p)} = A_0 \frac{K_A(p)\dfrac{1}{P_B(p)}}{1 + \dfrac{K_A(p)}{P_B(p)}} = \frac{A_0 K_A(p)}{P_B(p) + K_A(p)}.$$

Soit, dans le cas général,

$$K_A(p) = \frac{Q_A(p)}{P_A(p)}, \tag{3.100}$$

où le degré du polynôme $Q_A(p)$ est inférieur à celui du polynôme $P_A(p)$ égal à $n - m$. Par suite

$$\overline{X}(p) = \frac{A_0 Q_A(p)}{P_A(p) P_B(p) + Q_A(p)}. \tag{3.101}$$

Les conditions initiales de $x_d(t)$ ne sont invariables, c'est-à-dire indépendantes des paramètres de la fonction de transfert $K_A(p)$, que lorsque pour $t = 0$ la fonction $X(t)$ et ses $n - 1$ dérivées ne dépendent pas de ces paramètres et sont égales à zéro. La valeur initiale de l' i-ième dérivée étant assujettie à la transformée *) $\overline{X}(p)$, pour $\left(\frac{d^{i-1}X}{dt^{i-1}}\right)_{t=0} = 0$, par l'égalité

$$\left(\frac{d^i X}{dt^i}\right)_{t=0} = \lim_{p\to\infty} p^i \overline{X}(p) = \lim_{p\to\infty} \frac{A_0 p^i Q_A(p)}{P_A(p) P_B(p) + Q_A(p)}, \tag{3.102}$$

la condition

$$\left(\frac{d^i X}{dt^i}\right)_{t=0} = 0 \quad (i = 1, 2, \ldots, n-1) \tag{3.103}$$

signifie que le degré de $Q_A(p)$ doit être égal à zéro, c'est-à-dire $Q_A(p)$ doit être une grandeur constante. Ainsi, sans porter atteinte à la généralité, posons pour ce cas $Q_A(p) = 1$ et

$$K_A(p) = \frac{1}{P_A(p)} = \frac{1}{a_{n-m}p^{n-m} + \ldots + a_1 p + a_0}. \tag{3.104}$$

En vertu de (3.101), l'équation du système sous la forme des transformées s'écrit

$$[P_A(p) P_B(p) + 1] \overline{X}(p) = A_0, \tag{3.105}$$

et l'équation caractéristique, la même pour $X(t)$ et pour l'erreur transitoire $x_d(t)$, devient

$$P_A(p) P_B(p) + 1 = 0. \tag{3.106}$$

Par ailleurs, (3.98) donne l'équation caractéristique du système optimal sous la forme

$$H(p) = 0, \tag{3.107}$$

où

$$H(p) = \vartheta_n p^n + \vartheta_{n-1} p^{n-1} + \ldots + \vartheta_0. \tag{3.108}$$

La comparaison de (3.106) et de (3.107) conduit à l'égalité

$$\alpha H(p) = P_A(p) P_B(p) + 1, \tag{3.109}$$

*) La transformée utilisée ici est celle de Carson.

où α est un nombre quelconque différent de zéro. Si l'on choisit maintenant le polynôme $P_A(p)$ tel que l'égalité (3.109) soit réalisée, le système dans son ensemble devient optimal au sens indiqué ci-dessus.

En égalant les coefficients des mêmes puissances des premier et second membres de (3.109) on peut obtenir $n+1$ équations. Les inconnues de ces équations sont $\alpha, a_0, a_1, \ldots, a_{n-m}$, donc le nombre total des inconnues est $n-m+2$. Le système s'avère compatible et défini lorsque $n+1=n-m+2$, c'est-à-dire lorsque $m=1$. Ainsi, dans le cas considéré, l'objet ne peut être que de premier ordre.

Soit, par exemple, l'objet B qui représente le circuit intégrateur de fonction de transfert

$$K_B(p)=\frac{\xi_0}{p}=\frac{1}{\left(\frac{p}{\xi_0}\right)}=\frac{1}{P_B(p)}. \tag{3.110}$$

Supposons également que l'égalité (3.107) s'écrive

$$H(p)=p^2+2d_0\omega_0 p+\omega_0^2=0, \tag{3.111}$$

c'est-à-dire que tout le système doive constituer un circuit oscillant minimisant l'intégrale (cf. chapitre II)

$$I_V=\int_0^\infty\left\{x_d^2+2\sigma T_0^2\left(\frac{dx_d}{dt}\right)^2+T_0^4\left(\frac{d^2x_d}{dt^2}\right)^2\right\}dt, \tag{3.112}$$

où

$$T_0=\frac{1}{\omega_0}, \qquad d_0=\sqrt{\frac{1+\sigma}{2}}. \tag{3.113}$$

La fonction de transfert de l'organe de commande n'est alors que du premier degré

$$P_A(p)=a_1p+a_0. \tag{3.114}$$

En substituant dans (3.110) les expressions de $P_A(p)$, $P_B(p)$ et $H(p)$ données par (3.110), (3.111) et (3.114) on arrive à

$$\alpha(p^2+2d_0\omega_0p+\omega_0^2)=(a_1p+a_0)\left(\frac{p}{\xi_0}\right)+1, \tag{3.115}$$

d'où l'on tire trois équations pour les inconnues α, a_1, a_0

$$\alpha=\frac{a_1}{\xi_0}, \quad 2\alpha d_0\omega_0=\frac{a_0}{\xi_0}, \quad \alpha\omega_0^2=1.$$

Par conséquent

$$\alpha=\frac{1}{\omega_0^2}, \quad a_1=\frac{\xi_0}{\omega_0^2}, \quad a_0=\frac{2\xi_0d_0}{\omega_0},$$

et l'organe de commande optimal se trouve être un circuit inertie de fonction de transfert

$$P_A(p)=\frac{1}{a_1p+a_0}=\frac{1}{\frac{\xi_0}{\omega_0^2}p+\frac{2\xi_0d_0}{\omega_0}}=\frac{K_A}{1+pT_A}, \tag{3.116}$$

où

$$K_A = \frac{\omega_0}{2\xi_0 d_0}, \quad T_A = \frac{1}{2\omega_0 d_0}. \tag{3.117}$$

Dans ce problème on peut avoir ω_0 aussi grand que l'on veut, c'est-à-dire la durée du processus transitoire dans le système peut être aussi petite que l'on veut. Mais d'après (3.117) la grandeur K_A doit être choisie suffisamment grande, ce qui entraîne que dans ce cas la commande $u(t)$ devient également aussi grande que l'on veut. Pour restreindre cette grandeur, ne serait-ce que son intégrale, on peut ajouter la contrainte

$$I_u = \int_0^\infty u^2(t)\,dt \leqslant M.$$

Si l'on considère le régime transitoire le plus « pénible », on peut remplacer l'inégalité par l'égalité

$$I_u = \int_0^\infty u^2(t)\,dt = M. \tag{3.118}$$

L'équation de l'objet étant

$$\frac{dX}{dt} = \xi_0 u, \tag{3.119}$$

il résulte pour l'exemple concerné, avec $x_s = 0$ et $X_0 = \text{const}$,

$$\frac{dx_d}{dt} = \frac{d}{dt}(x - x_s) = \frac{dx}{dt} = \frac{d}{dt}(X_0 - X) = \frac{d}{dt}(A_0 - X) = -\frac{dX}{dt} = -\xi_0 u$$

et la condition (3.119) devient égalité

$$I_u = \int_0^\infty u^2(t)\,dt = \int_0^\infty \frac{1}{\xi_0^2}\left(\frac{dx_d}{dt}\right)^2 dt = M. \tag{3.120}$$

Introduisons le multiplicateur de Lagrange λ. Si la condition à observer est (3.120), la minimisation doit porter sur l'intégrale

$$Q = I_v + \lambda I_u =$$

$$= \int_0^\infty \left\{x_d^2 + \left(2\sigma T_0^2 + \frac{\lambda}{\xi_0^2}\right)\left(\frac{dx_d}{dt}\right)^2 + T_0^4\left(\frac{d^2x_d}{dt^2}\right)^2\right\} dt. \tag{3.121}$$

Posons

$$2\sigma' T_0^2 = 2\sigma T_0^2 + \frac{\lambda}{\xi_0^2}.$$

On obtient alors pour l'intégrale Q la même expression que pour l'intégrale I_v, seulement σ est remplacé par σ'. On en déduit que la solution a également la même forme, seulement d_0 est remplacé

par d_0':

$$d_0' = \sqrt{\frac{\sigma'+1}{2}} = \sqrt{\frac{1}{2}\left(\sigma+1+\frac{\lambda}{2T_0^2\xi_0^2}\right)} =$$

$$= \sqrt{\frac{1}{2}\left(2d_0^2+\frac{\lambda\omega_0^2}{2\xi_0^2}\right)} = \sqrt{d_0^2+\frac{\lambda\omega_0^2}{4\xi_0^2}}. \qquad (3.122)$$

D'autre part, le paramètre λ doit être choisi tel qu'il vérifie la condition (3.120). Cette condition, où figure également la valeur A_0 de la consigne, donne encore une relation entre les paramètres, alors que les deux autres sont données par les équations du type (3.117)

$$k_A = \frac{\omega_0}{2\xi_0 d_0'}, \quad T_A = \frac{1}{2\omega_0 d_0'}.$$

Nous ne parviendrons plus à rendre le processus transitoire $x_d(t)$ susceptible de minimiser l'intégrale (3.112). La trajectoire optimale $x_d(t)$ satisfait à l'équation

$$\frac{d^2x_d}{dt^2} + 2d_0'\omega_0\frac{dx_d}{dt} + \omega_0^2 x_d = 0. \qquad (3.123)$$

Calculons pour cette courbe l'intégrale

$$I = \int_0^\infty \left(\frac{dx_d}{dt}\right)^2 dt = \int_0^\infty U\,dt \qquad (3.124)$$

d'après la méthode exposée dans [3.25]. Les conditions initiales de x_d s'écrivent (avec $x_s = 0$ et $x = x_d$)

$$(x_d)_{t=0} = (X_0 - X)_{t=0} = A_0,$$

$$\left(\frac{dx_d}{dt}\right)_{t=0} = \left[\frac{d}{dt}(X_0) - \frac{d}{dt}(X)\right]_{t=0} = 0.$$

Il vient

$$I = B_{11}A_0^2, \qquad (3.125)$$

c'est-à-dire la grandeur I est égale à la valeur initiale de la forme quadratique

$$W = B_{11}x_1^2 + 2B_{12}x_1x_2 + B_{22}x_2^2, \qquad (3.126)$$

qui vérifie la condition

$$U = -\frac{dW}{dt}. \qquad (3.127)$$

On a posé ici

$$x_1 = x_d, \quad x_2 = \frac{dx_d}{dt}.$$

On en déduit

$$U = \left(\frac{dx_d}{dt}\right)^2 = x_2^2. \qquad (3.128)$$

Au lieu de l'équation (3.123) nous pouvons écrire deux équations de premier ordre

$$\left.\begin{aligned} \frac{dx_1}{dt} &= x_2, \\ \frac{dx_2}{dt} &= -\omega_0^2 x_1 - 2d_0'\omega_0 x_2. \end{aligned}\right\} \tag{3.129}$$

Ecrivons l'égalité (3.127) sous une forme développée

$$-U = \frac{\partial W}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial W}{\partial x_2}\frac{dx_2}{dt}.$$

Portant ici les expressions (3.128), (3.126) et (3.129) on aboutit à l'équation

$$\begin{aligned} -x_2^2 = 2B_{11}x_1x_2 + 2B_{12}x_2^2 - 2B_{12}\omega_0^2x_1^2 - 2B_2\omega_0^2x_1x_2 - \\ -4B_{12}d_0'\omega_0x_1x_2 - 4d_0'\omega_0B_{22}x_2^2. \end{aligned} \tag{3.130}$$

En égalant les coefficients des termes identiques des premier et second membres, on obtient trois équations

$$\left.\begin{aligned} -2B_{12}\omega_0^2 = 0;\ \ 2B_{11} - 2B_{22}\omega_0^2 - 4B_{12}d_0'\omega_0 = 0; \\ 2B_{12} - 4d_0'\omega_0B_{22} = -1, \end{aligned}\right\} \tag{3.131}$$

d'où l'on trouve

$$B_{12} = 0,\ \ B_{22} = \frac{1}{4d_0'\omega_0},\ \ B_{11} = \frac{\omega_0}{4d_0'}. \tag{3.132}$$

(3.125) devient

$$I = B_{11}A_0^2 = \frac{\omega_0 A_0^2}{4d_0'}. \tag{3.133}$$

Compte tenu de (3.133), la condition (3.121) donne

$$\frac{\omega_0 A_0^2}{4d_0'} = \xi_0^2 M, \tag{3.134}$$

d'où

$$d_0' = \frac{A_0^2\omega_0}{4\xi_0^2 M} = \varkappa\omega_0, \tag{3.135}$$

où

$$\varkappa = \frac{A_0^2}{4\xi_0^2 M}. \tag{3.136}$$

Connaissant d_0', on peut trouver λ de (3.122) mais nous n'allons pas nous y attarder.

La substitution de (3.135) dans les expressions de k_A et de T_A donne les paramètres de l'organe de commande optimal A

$$k_A = \frac{2\xi_0 M}{A_0^2},\ \ T_A = \frac{2\xi_0^2 M}{\omega_0^2 A_0^2}. \tag{3.137}$$

On en tire qu'avec un seuil admissible M peu grand et une grande valeur de A_0 les valeurs du gain k_A et de la constante de temps T_A

doivent être petites. On voit aisément que la solution de l'équation (3.123) de la trajectoire optimale n'est plus un processus transitoire aussi rapide que l'on veut, quelque grande que soit ω_0. En effet, l'équation caractéristique de la trajectoire optimale s'écrit

$$p^2 + 2d_0'\omega_0 p + \omega_0^2 = 0. \tag{3.138}$$

Portons la valeur de d_0' tirée de (3.135). L'équation (3.137) se met alors sous la forme

$$p^2 + 2\varkappa\omega_0^2 p + \omega_0^2 = 0. \tag{3.139}$$

Ses racines sont

$$p_{1,2} = -\varkappa\omega_0^2 \pm \sqrt{\varkappa^2\omega_0^4 - \omega_0^2} = \omega_0^2\left(-\varkappa \pm \sqrt{\varkappa^2 - \frac{1}{\omega_0^2}}\right). \tag{3.140}$$

Avec $\omega_0^2 \to \infty$ l'une des racines

$$p_1 \cong \omega_0^2\left(-\varkappa - \sqrt{\varkappa^2}\right) = -2\varkappa\omega_0^2 \tag{3.141}$$

tend vers $-\infty$ et donc n'influe plus sur le processus transitoire. La deuxième racine

$$p_2 = \omega_0^2\left(-\varkappa + \varkappa\sqrt{1 - \frac{1}{\varkappa^2\omega_0^2}}\right) =$$

$$= \omega_0^2\varkappa\left(-1 + 1 - \frac{1}{2\varkappa^2\omega_0^2} + o\left(\frac{1}{\omega_0^2}\right)\right) \to -\frac{1}{2\varkappa} \tag{3.142}$$

avec $\omega_0^2 \to \infty$ ne dépend plus de ω_0^2. La figure 3.18 montre la tendance des racines p_1 et p_2 pour $\omega_0 \to \infty$. Ainsi le processus transitoire devient l'exponentielle $A_0 \exp\{p_2 t\} = A_0 \exp\left\{-\frac{1}{2\varkappa}\right\}$ qui ne dépend pas de ω_0^2, c'est-à-dire le processus optimal ne peut pas être aussi rapide que l'on veut. On peut trouver des exemples analogues de synthèse des systèmes optimaux ou de recherche des paramètres optimaux dans [3.15], [1.5], [1.9], [3.28].

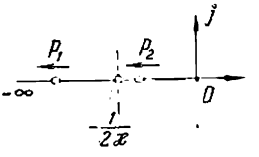


Fig. 3.18.

Les méthodes classiques du calcul des variations permettent également de résoudre sans peine les problèmes de synthèse dans le cas où l'équation de l'objet est d'un ordre plus élevé, si seulement cette dernière est linéaire à coefficients constants, alors que le critère d'optimalité et les expressions à restreindre par les égalités sont des intégrales des formes quadratiques.

La résolution devient plus difficile lorsqu'on ajoute des contraintes sous la forme d'inégalités; de plus, comme nous l'avons dit au chapitre II, les difficultés accroissent lorsque le résultat doit être donné par des fonctions discontinues. Pour éviter ces difficultés et simplifier la procédure de résolution par les méthodes classiques,

on remplace parfois le problème initial par un problème approché, qui diffère du problème initial par ce que ses conditions comportent des fonctions aussi « lisses » qu'il faut pour obtenir la solution de même suffisamment lisse. Considérons à titre d'exemple le problème examiné dans [3.29]. Soient les équations de l'objet de commande

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r) \quad (i = 1, 2, \ldots, n); \tag{3.143}$$

la tâche consiste à choisir les commandes $u_1, u_2, \ldots, u_r$ telles qu'elles minimisent l'intégrale

$$Q = \int_0^T V(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r)\, dt, \tag{3.144}$$

où T n'est pas fixé, alors que la fonction V est supposée dérivable par rapport à tous ses arguments, de même que les fonctions f_i $(i = 1, 2, \ldots, n)$. Les valeurs initiale et finale du vecteur $\bar{x} = (x_1, x_2, \ldots, x_n)$ sont également données.

Soit la contrainte supplémentaire

$$\psi(\bar{u}) = \psi(u_1, u_2, \ldots, u_r) \leqslant 0. \tag{3.145}$$

Pour approcher cette contrainte on peut introduire la «fonction de pénalisation » en imposant une « pénalité » très élevée pour l'inobservation de la condition (3.145). Soit la valeur de la « pénalité »

$$L(\bar{u}) = \begin{cases} 0, & \psi \leqslant 0, \\ k\psi^2, & \psi > 0, \end{cases} \tag{3.146}$$

où $k > 0$ est suffisamment grand. L'inobservation même négligeable de la condition (3.145) donne lieu à une « pénalité » d'une valeur importante. Le problème se formule alors comme suit : la commande idéale doit assurer le minimum de l'intégrale (3.144) tout en observant la condition

$$\int_0^T L(\bar{u})\, dt = 0. \tag{3.147}$$

La condition (3.147) est exacte, car si elle n'est pas violée, ψ ne devient positif en aucun instant (sauf, peut-être, dans le cas de l'ensemble de mesure nulle, non considéré ici). L'artifice qui consiste à remplacer la condition (3.145) par l'égalité (3.147) où figure l'intégrale peut simplifier sensiblement la résolution du problème (cf. également [3.27]). Considérons maintenant au lieu de l'intégrale initiale (3.144) une nouvelle intégrale

$$Q' = Q + \int_0^T L(\bar{u})\, dt = \int_0^T V\, dt + \int_0^T L\, dt. \tag{3.148}$$

Pour être plus précis, il fallait introduire un multiplicateur de Lagrange devant le deuxième terme. Pourtant lorsque le coefficiene k de la formule (3.146) est suffisamment grand, même une faible « pénétration » de $\bar{u}$ dans le domaine interdit $\psi > 0$ provoque un si grand accroissement du deuxième terme que Q' s'écarte fortement du minimum. Ainsi le point $\bar{u}$ associé au minimum Q' se situe soit à l'intérieur du domaine $\psi \leqslant 0$, soit pratiquement sur la frontière de ce domaine. Nous pouvons donc nous guider maintenant par le seul critère (3.148) sans faire appel à (3.147).

Les équations de l'objet (3.143) jouent dans le cas considéré le rôle des contraintes sous la forme d'égalités

$$\dot{x}_i - f_i(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_r) = 0 \qquad (i = 1, 2, \ldots, n). \tag{3.149}$$

Pour en tenir compte, il faut introduire les multiplicateurs de Lagrange λ_i [cf. chapitre II, équations (2.153) et (2.154)] et au lieu de Q' minimiser une autre grandeur Q'',

$$Q'' = \int_0^T V\,dt + \int_0^T L\,dt + \\ + \int_0^T \left\{ \sum_{i=1}^n \lambda_i [\dot{x}_i - f_i(x_1, \ldots, x_n, u_1, \ldots, u_r)] \right\} dt = \int_0^T F^*\,dt, \tag{3.150}$$

où

$$F^* = V(x_1, \ldots, x_n) + L(u_1, \ldots, u_r) + \\ + \sum_{j=1}^n \lambda_j [\dot{x}_j - f_j(x_1, \ldots, x_n, u_1, \ldots, u_r)]. \tag{3.151}$$

Maintenant il faut composer les équations d'Euler-Lagrange en considérant x_i et u_j comme variables ordinaires :

$$\frac{\partial F^*}{\partial x_i} - \frac{d}{dt}\frac{\partial F^*}{\partial \dot{x}_i} = 0 \quad (i = 1, 2, \ldots, n) \tag{3.152}$$

et

$$\frac{\partial F^*}{\partial u_k} - \frac{d}{dt}\frac{\partial F^*}{\partial \dot{u}_k} = 0 \quad (k = 1, 2, \ldots, r). \tag{3.153}$$

Par ailleurs, il faut encore remplir n conditions (3.149). Au total on obtient $2n + r$ équations pour les inconnues x_i $(i = 1, 2, \ldots, n)$, λ_j $(j = 1, 2, \ldots, n)$ et u_k $(k = 1, 2, \ldots, r)$, dont le nombre total est également $2n + r$. En portant dans (3.152) et

(3.153) l'expression (3.151) on aboutit aux équations

$$\frac{\partial V}{\partial x_i} - \sum_{j=1}^{n} \lambda_j \frac{\partial f_j}{\partial x_i} - \frac{d\lambda_i}{dt} = 0 \quad (i = 1, 2, \ldots, n) \tag{3.154}$$

et

$$\frac{\partial L}{\partial u_k} - \sum_{j=1}^{n} \lambda_j \frac{\partial f_j}{\partial u_k} = 0 \quad (k = 1, 2, \ldots, r). \tag{3.155}$$

Si on exprime les u_k de (3.155) à l'aide des autres variables du système $\left(\frac{du_k}{dt}\right.$ ne font nulle part partie des égalités$\left.\right)$, il reste alors $2n$ équations aux $2n$ inconnues $x_1, x_2, \ldots, x_n$ et $\lambda_1, \lambda_2, \ldots, \lambda_n$, alors que les conditions aux limites données sont au nombre de $2n$.

Traitons l'exemple d'un objet linéaire. Soient ses équations qui s'écrivent

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij} x_j + \delta_{i1} u, \tag{3.156}$$

où

$$\delta_{i1} = \begin{cases} 0, & i \neq 1, \\ 1, & i = 1. \end{cases} \tag{3.157}$$

Ainsi, u n'est présent que dans la première des équations (3.156).

Supposons qu'il faut minimiser la durée du processus transitoire. On a

$$Q = \int_0^T 1 \cdot dt = T. \tag{3.158}$$

Par conséquent $V = 1$. Ensuite, soit l'unique commande $u(t)$ restreinte par la condition

$$|u| \leqslant M. \tag{3.159}$$

On peut donc poser

$$L(u) = L_1(u) + L_2(u), \tag{3.160}$$

où

$$\left.\begin{aligned} L_1(u) &= \begin{cases} 0, & u + M > 0, \\ k(u+M)^2, & u + M < 0. \end{cases} \\ L_2(u) &= \begin{cases} 0, & M - u > 0, \\ k(M-u)^2, & M - u < 0. \end{cases} \end{aligned}\right\} \tag{3.161}$$

Ces fonctions sont représentées sur la figure 3.19,*a*. D'après (3.151) la fonction F^* pour l'exemple traité est de la forme

$$F^* = 1 + [L_1(u) + L_2(u)] + \sum_{k=1}^{n} \lambda_k \left(\dot{x}_k - \sum_{j=1}^{n} a_{kj} - x_j - \delta_{k1} u\right). \quad (3.162)$$

Le système des équations d'Euler (3.154) et (3.155) s'écrit alors

$$\frac{d\lambda_i}{dt} = -\sum_{j=1}^{n} a_{ji}\lambda_j \quad (i = 1, 2, \ldots, n), \quad (3.163)$$

$$\lambda_1 = \varphi(u) = \varphi_1(u) + \varphi_2(u), \quad (3.164)$$

où

$$\varphi(u) = \frac{dL}{du}$$

et

$$\left.\begin{aligned} \varphi_1(u) &= \begin{cases} 0, & u + M > 0, \\ 2k(u + M), & u + M < 0, \end{cases} \\ \varphi_2(u) &= \begin{cases} 0, & M - u > 0, \\ -2k(M - u), & M - u < 0. \end{cases} \end{aligned}\right\} \quad (3.165)$$

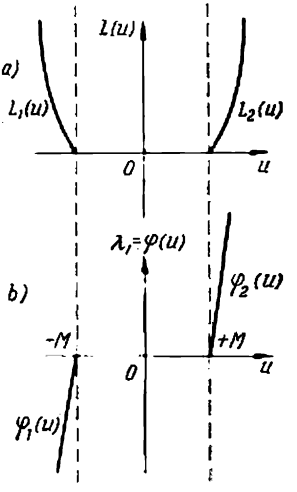


Fig. 3.19.

La figure 3.19,*b* représente la courbe $\varphi(u)$. Elle est composée de la portion $\varphi(u) = 0$ avec $|u| \leqslant M$ et des portions linéaires de pente k. Plus la valeur de k est grande, plus les droites inclinées sont proches des demi-droites verticales matérialisées par le trait interrompu.

Les inconnues λ_i vérifient le système d'équations (3.163) adjoint au système d'équations (3.156) de l'objet. Si les racines de l'équation caractéristique associée à (3.156) sont négatives et réelles, on peut montrer que les racines de l'équation caractéristique correspondant au système (3.163) adjoint sont également négatives et réelles. Soient ces racines égales à $\gamma_1, \gamma_2, \ldots, \gamma_n$. On a dans le cas général

$$\lambda_1 = C_1 e^{\gamma_1 t} + \ldots + C_n e^{\gamma_n t}. \quad (3.166)$$

On sait que cette fonction change de signe pas plus de $n - 1$ fois et compte, par conséquent, dans le cas général, n intervalles de signe constant. Mais l'égalité (3.164) permet d'obtenir à tout instant la grandeur $u(t)$ associée à la valeur $\lambda_1(t)$. La figure 3.19 montre que $\lambda_1 = \varphi(u)$ ne peut changer de signe que lorsque u passe par saut du domaine $u < -M$ dans celui où $u > +M$, ou inversement. Lorsque la pente k des portions linéaires est suffisamment grande, les valeurs correspondantes de u sont égales à $+M$ et $-M$. De cette

façon, le nombre total des intervalles de constance de $u = \pm M$ ne dépasse pas n. Nous avons démontré ainsi le théorème des n intervalles.

Un exemple intéressant de synthèse d'un système de commande optimal d'une pile chimique est donné dans [3.32] (cf. également

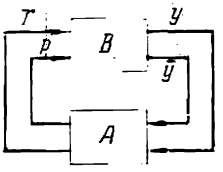


Fig. 3.20.

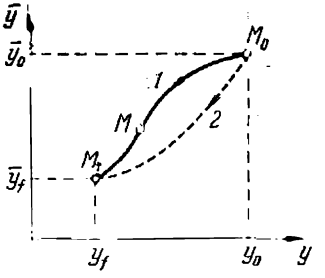


Fig. 3.21.

[3.33]). Le problème consiste à minimiser la durée de la réaction, ce qui correspond à la productivité maximale de la pile.

Soit la pile B (fig. 3.20) qui contient le mélange de trois composantes, dont les concentrations relatives sont notées y, $\bar{y}$ et z. La somme des concentrations relatives est égale à l'unité ; par conséquent

$$y + \bar{y} + z = 1. \tag{3.167}$$

Au cours des réactions chimiques dans B ces concentrations varient des valeurs initiales y_0, $\bar{y}_0$, z_0 jusqu'aux valeurs finales y_f, $\bar{y}_f$, z_f. Le temps nécessaire pour atteindre ces valeurs dépend des lois qui régissent pendant la réaction la température T et la pression P dans la pile. Il faut concevoir un organe de commande A susceptible de réaliser une modification optimale des paramètres de commande T et P pour rendre le processus le plus rapide possible.

Comme z peut être calculé à partir de l'équation (3.167) d'après les valeurs données de y et de $\bar{y}$, on peut considérer seulement deux composantes indépendantes du mélange y et $\bar{y}$. Traçons le plan de phase de coordonnées cartésiennes y et $\bar{y}$ (fig. 3.21). L'état initial du mélange correspond au point M_0 de coordonnées y_0 et $\bar{y}_0$ et l'état final au point M_f de coordonnées y_f et $\bar{y}_f$. Le point représentatif M qui part de M_0 peut arriver en M_f suivant différentes trajectoires, par exemple, *1* ou *2*, définies par les lois de variation de T et de P. Il faut trouver une loi (et par conséquent une trajectoire du point M dans le plan de phase) telle qu'elle minimise le temps nécessaire pour aller de M_0 en M_f.

Dans le cas considéré les constantes de temps de l'organe de commande A sont négligeables par rapport au temps de la réaction. Les équations de la trajectoire de l'objet sont celles de la cinétique de la pile composées sur la base de la loi d'action de masse. Cette loi dit que la vitesse de la réaction est proportionnelle aux concentrations des substances réagissantes. Dans notre exemple les substances dont les concentrations sont désignées y, $\bar{y}$ et z réagissent avec l'hydrogène refoulé dans la pile B. Si l'on considère que la quantité d'hydrogène fournie à la pile par unité de temps est constante, la vitesse de diminution d'une composante quelconque du mélange est proportionnelle à sa quantité. Introduisons le coefficient de proportionnalité k affecté d'un indice convenable.

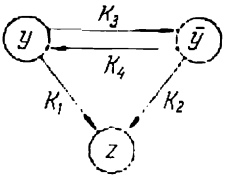


Fig. 3.22.

Les réactions sont schématisées par la figure 3.22. La direction des réactions est indiquée par les flèches munies d'un coefficient de proportionnalité. Le schéma montre que la substance y se transforme en $\bar{y}$ avec le coefficient de proportionnalité k_3 et en z avec k_1. A son tour la substance $\bar{y}$ se transforme en y avec le coefficient de proportionnalité k_4 et en z avec k_2. On peut donc écrire ce qu'on appelle d'habitude les équations cinétiques de l'objet B :

$$\frac{dy}{dt} = (-k_1 + k_3)\, y + k_4 \bar{y} \tag{3.168}$$

et

$$\frac{d\bar{y}}{dt} = (-k_2 + k_4)\, \bar{y} + k_3 y. \tag{3.169}$$

Les coefficients k_i dépendent de la température T et de la pression P. La relation caractéristique des réactions chimiques s'écrit

$$k_i = a_i P^{n_i} \exp\left\{-\frac{b_i}{T}\right\}, \tag{3.170}$$

où a_i, n_i, b_i sont des constantes. Si pour simplifier le problème on admet la température T constante, ce qui au fond n'est pas la meilleure façon de conduire la réaction, on a

$$k_i = A_i P^{n_i}, \tag{3.171}$$

où A_i sont des constantes ($i = 1, 2, 3, 4$). Le problème consiste à définir la fonction optimale $P(t)$ telle qu'elle minimise la durée de passage du système du point M_0 en M_f (cf. fig. 3.21).

Pour rendre possible la formulation du problème variationnel correspondant, il est avantageux de passer à de nouvelles variables

$$w = \frac{\bar{y}}{y} \tag{3.172}$$

et

$$q = \ln \frac{y_0}{y}. \tag{3.173}$$

A l'instant initial $q = \ln 1 = 0$. A mesure que le temps t croît la grandeur y diminue, car diminue la concentration de la substance initiale. Par conséquent q croît monotonément avec le temps. Pour simplifier les équations nous pouvons considérer q comme une nouvelle variable indépendante.

Pour passer à de nouvelles variables, transformons d'abord l'équation (3.168)

$$\frac{1}{y}\frac{dy}{dt} = -(k_1 + k_3) + k_4 \frac{\bar{y}}{y}. \tag{3.174}$$

Compte tenu de (3.172) on a

$$\frac{d}{dt}(\ln y) = -(k_1 + k_3) + k_4 w. \tag{3.175}$$

Etant donné que $q = \ln y_0 - \ln y$, la dernière équation se met sous la forme

$$\frac{dq}{dt} = (k_1 + k_3) - k_4 w. \tag{3.176}$$

Développons ensuite l'expression de $\frac{dw}{dt}$:

$$\frac{dw}{dt} = \frac{d}{dt}\left(\frac{\bar{y}}{y}\right) = \frac{y\frac{d\bar{y}}{dt} - \bar{y}\frac{dy}{dt}}{y^2} = \frac{1}{y}\frac{d\bar{y}}{dt} - \frac{\bar{y}}{y}\left[\frac{1}{y}\frac{dy}{dt}\right] =$$

$$= \frac{1}{y}\frac{d\bar{y}}{dt} - w\,[-(k_1 + k_3) + k_4 w]. \tag{3.177}$$

L'expression entre crochets est remplacée ici en conformité de (3.174). L'équation (3.169) entraîne

$$\frac{1}{y}\frac{d\bar{y}}{dt} = k_3 - (k_2 + k_4)\, w. \tag{3.178}$$

En portant cette expression dans (3.177) on passe à l'équation

$$\frac{dw}{dt} = k_3 + (k_1 + k_3 - k_2 - k_4)\, w - k_4 w^2. \tag{3.179}$$

Après avoir divisé terme à terme les équations (3.179) et (3.176) on obtient

$$\frac{dw}{dq} = \frac{k_3 + (k_1 + k_3 - k_2 - k_4)\, w - k_4 w^2}{k_1 + k_3 - k_4 w}. \tag{3.180}$$

L'équation (3.176) entraîne

$$dt = \frac{dq}{k_1 + k_3 - k_4 w}. \tag{3.181}$$

On peut maintenant calculer la durée de la réaction t_f, car à l'instant initial $q=0$ et à l'instant final $q=q_f=\ln\left(\frac{y_0}{y_f}\right)$:

$$t_f=\int_0^{q_f}\frac{dq}{k_1+k_3-k_4w}. \tag{3.182}$$

Le problème se ramène ainsi à la recherche d'une trajectoire dans le plan $(q,\ w)$ et d'une loi de variation de $P\ (q)$ telles qu'elles minimisent l'intégrale (3.182).

Ces équations peuvent être simplifiées, car pour la réaction examinée en [3.32] ont lieu les relations

$$\frac{k_2}{k_1}\cong\frac{k_4}{k_3},\quad n_1=n_2,\quad n_3=n_4. \tag{3.183}$$

De plus, $\frac{k_1}{k_3}<1$. Posons

$$\left.\begin{aligned} g=\frac{k_1}{k_3}=\left(\frac{A_3}{A_5}\right)P^{(n_1-n_3)},\quad A=\frac{A_1^{m-1}}{A_3^{m-1}},\\ B=\frac{k_4}{k_3}=\frac{A_4}{A_3},\quad C=\frac{k_2}{k_3}=\frac{A_2}{A_3},\quad m=\frac{n_1}{n_1-n_3}.\end{aligned}\right\} \tag{3.184}$$

Il vient alors

$$P=\left[\frac{g^m}{AA_3}\right]^{1/n_3}. \tag{3.185}$$

Tous les coefficients k_i peuvent être considérés comme fonctions de la grandeur g, qui à son tour est assujettie à q. L'intégrale (3.182) peut s'écrire

$$t_f=\int_0^{q_f}F\ (w,\ g)\ dq. \tag{3.186}$$

La grandeur w est liée à q par l'équation (3.180) que nous allons mettre sous la forme

$$\varphi\ (w,\ g,\ w')=f\ (w,\ g)-w'=0. \tag{3.187}$$

Ici $w'=\frac{dw}{dq}$. Dans ce qui suit la dérivation par rapport à q est notée par un apostrophe.

Ainsi il faut rechercher les fonctions $g\ (q)$ et $w\ (q)$ qui minimisent l'intégrale (3.186) sous la condition supplémentaire (3.187). Pour trouver l'équation d'Euler composons la fonction auxiliaire

$$F^*=F+\lambda\ (q)\ \varphi=F\ (w,\ g)+\lambda\ (q)\ [f\ (w,\ g)-w'], \tag{3.188}$$

$\lambda(q)$ étant le multiplicateur de Lagrange. Pour déterminer les trois fonctions λ, w et g on dispose de deux équations d'Euler

$$\frac{\partial F^*}{\partial w} - \frac{d}{dq}\left(\frac{\partial F^*}{\partial w'}\right) = 0, \tag{3.189}$$

$$\frac{\partial F^*}{\partial g} - \frac{d}{dq}\left(\frac{\partial F^*}{\partial g'}\right) = 0 \tag{3.190}$$

et de l'équation de liaison (3.187). En portant dans (3.189) et (3.190) les valeurs des dérivées par rapport à F^* tirées de (3.188), il vient

$$\lambda(q)\frac{\partial f(w, g)}{\partial w} + \frac{\partial F(w, g)}{\partial w} + \frac{\partial \lambda(q)}{\partial q} = 0 \tag{3.191}$$

et

$$\lambda(q)\frac{\partial f(w, g)}{\partial w} + \frac{\partial F(w, g)}{\partial g} = 0. \tag{3.192}$$

Dans les deux dernières équations on peut éliminer la fonction $\lambda(q)$. (3.191) donne

$$\lambda(q) = -\frac{\dfrac{\partial F(w, g)}{\partial g}}{\dfrac{\partial f(w, g)}{\partial g}} = \psi(w, g), \tag{3.193}$$

donc

$$\frac{d\lambda(q)}{dq} = \frac{\partial \psi}{\partial w}\frac{\partial w}{\partial q} + \frac{\partial \psi}{\partial g}\frac{dg}{dq} = \frac{\partial \psi}{\partial w} f(w, g) + \frac{\partial \varphi}{\partial g} g'. \tag{3.194}$$

En portant (3.193) et (3.194) dans (3.191) on obtient l'équation

$$\psi(w, g)\frac{\partial f(w, g)}{\partial w} + \frac{\partial F(w, g)}{\partial w} + \frac{\partial \psi(w, g)}{\partial w}\cdot f(w, g) + \frac{\partial \psi(w, g)}{\partial g} g' = 0, \tag{3.195}$$

qui est une équation différentielle du premier ordre par rapport à g. Après substitution des valeurs des dérivées cette équation devient

$$g' = (1 + w - w')\left[\frac{g}{(m-1)(1-Bw)} - \frac{1-Bw}{m}\right]. \tag{3.196}$$

Les notations de (3.184) permettent d'écrire les équations (3.180) et (3.181) sous la forme

$$w' = w + 1 - \frac{g(1+Cw)}{g+1-Bw} \tag{3.197}$$

et

$$t' = \frac{Ag^{1-m}}{g+1-Bw}. \tag{3.198}$$

La résolution simultanée des équations (3.196) et (3.197) donne les courbes cherchées $g(q)$ et $w(q)$. A l'instant considéré la valeur initiale $w(0)$ est connue. On connaît également la valeur finale $w_f = w(q_f)$, mais on ignore la valeur $g(0)$. Or, c'est précisément cette

valeur qu'il faut déterminer, car connaissant g à l'instant courant on peut calculer d'après la formule (3.185) la valeur nécessaire de la pression P. Pour obtenir $g(0)$ il faut résoudre le problème aux limites. Cette grandeur doit être choisie telle que le point représentatif se déplace dans le plan (q, w) de la position initiale $\left(q_0 = 0, w = w_0 = \frac{\bar{y}_0}{y_0}\right)$ à la position finale $\left(q = q_f = \ln\left(\frac{y_0}{y_f}\right), w = w_f = \frac{\bar{y}_f}{y_f}\right)$. La trajectoire doit vérifier les équations (3.196) et (3.197). La valeur t_f calculée simultanément d'après la formule (3.186) donne le temps nécessaire de la réaction.

Dans le prototype du calculateur utilisé pour résoudre ce problème, le choix de la valeur $g(0)$ était automatisé, alors que les valeurs courantes y_0 et $\bar{y}_0$, considérées comme valeurs initiales, étaient les données moyennes fournies par des appareils de mesure et introduites à la main. Le choix de $g(0)$ était guidé par une série de calculs approchés.

Cette procédure consiste en ce qui suit.

On intègre les équations (3.196) et (3.197) en fonction du « temps » q pour une valeur initiale w_0 donnée et un $g(0)$ quelconque et on mesure la valeur w'_f à l'instant q_f. Si $w_f - w'_f > 0$, c'est-à-dire si la grandeur w_f est trop petite, il faut augmenter $g(0)$. Cet accroissement de $g(0)$ a été rendu proportionnel à la différence mesurée $w_f - w'_f$. Après quelques cycles, la valeur de cette différence diminue pour devenir négligeable et de cette façon on définit automatiquement la valeur de $g(0)$ qui sert à calculer la valeur courante de la pression P, pression qu'il faut maintenir dans la pile à l'instant considéré.

Les exemples traités montrent que l'application des méthodes variationnelles classiques peut présenter et présente effectivement un intérêt dans certains domaines de la théorie des systèmes optimaux. Pourtant l'usage de nouveaux principes, comme la programmation dynamique et le principe du maximum, est plus riche de promesses.

§ 3. Application de la programmation dynamique

Les ouvrages [3.34, 3.35] traitent d'un objet linéaire ayant dans le cas général des paramètres variables dont la réponse impulsionnelle est notée $g(t, \tau)$. Cela signifie que la grandeur de sortie $x(t)$, observée à l'instant t, est liée à la grandeur d'entrée $u(\tau)$, $(-\infty < \tau \leqslant t)$ par la relation

$$x(t) = \int_{-\infty}^{t} g(t, \tau)\, u(\tau)\, d\tau. \tag{3.199}$$

Pour les systèmes linéaires à coefficients constants $g(t, \tau) = g(t - \tau)$ l'expression (3.199) devient intégrale de Duhamel usuelle. Dans ce qui suit, à chaque instant donné, l'état de l'objet est caractérisé par les valeurs de n coordonnées $x_1, x_2, \ldots, x_n$ ou par le vecteur $\overline{x}$.

Choisissons comme critère d'optimalité une certaine intégrale prise par rapport aux valeurs *futures* du temps dans l'intervalle de t à $t + T$

$$Q = \int_t^{t+T} \{\lambda(\sigma) f_x [\overline{X}(\sigma) - \overline{x}(\sigma)] + f_u [\overline{U}(\sigma) - \overline{u}(\sigma)]\} d\sigma. \quad (3.200)$$

Le problème consiste à choisir la valeur courante du vecteur $\overline{u}(t)$ telle qu'elle minimise la grandeur Q. L'expression (3.200) comporte sous le signe somme la fonction f_x de la différence $[\overline{X}(\sigma) - \overline{x}(\sigma)]$, où $\overline{X}(\sigma)$ est la fonction « idéale » donnée du vecteur $\overline{x}$; par ailleurs $f_x(0) = 0$. La fonction f_u de la différence $[\overline{U}(\sigma) - \overline{u}(\sigma)]$, où $\overline{U}(\sigma)$ est la commande « idéale », possède également la propriété $f_u(0) = 0$. Dans ce qui suit nous posons

$$f_x(v) = f_u(v) = v^2.$$

D'autre part, $\lambda(\sigma)$ est une certaine fonction de pondération qui donne le poids relatif des termes f_x et f_u. L'intervalle d'intégration est pris fini du fait que dans les problèmes pratiques on peut généralement établir la valeur finie de T, au-delà de laquelle la prévision de l'avenir ne présente plus aucun intérêt *).

Dans le cas d'une commande optimale, la grandeur Q acquiert une valeur minimale S qui ne dépend que de l'état de l'objet à l'instant courant t pris pour l'instant initial. Ainsi

$$S = S(x_1, x_2, \ldots, x_n, t) = S[\overline{x}(t), t] = \min_{u(\sigma) \in \Omega(u),\ t \leqslant \sigma \leqslant t+T} Q. \quad (3.201)$$

En remplaçant Q par l'expression (3.200), il vient

$$S[\overline{x}(t), t] = \min_{\substack{u(\sigma) \in \Omega(u) \\ t \leqslant \sigma \leqslant t+T}} \int_t^{t+T} \{\lambda(\sigma) f_x [\overline{X}(\sigma) - \overline{x}(\sigma)] + f_u [\overline{U}(\sigma) - \overline{u}(\sigma)]\} d\sigma. \quad (3.202)$$

Notons que

$$S[\overline{x}(t+T), t+T] = 0. \quad (3.203)$$

---

*) Notons qu'à l'exemple considéré est applicable le principe du maximum dans sa forme généralisée par A. Boutkovski aux objets régis par des équations intégrales (cf. [3.43, 3.44]).

En effet, d'après (3.200), la valeur de l'intégrale Q est égale à zero lorsque la limite d'intégration inférieure est égale à $t+T$. C'est pourquoi également $S = \min Q = 0$ si la valeur courante du temps coïncide avec l'instant final.

De même que dans le chapitre II, cherchons la condition d'optimalité nécessaire. A cet effet divisons l'intervalle de temps de t à $t+T$ en deux intervalles $[t \leqslant \sigma \leqslant t+\varepsilon]$ et $[t+\varepsilon < \sigma \leqslant t+T]$, où ε est petit. On a

$$S[\bar{x}(t), t] = \min_{\substack{u(\sigma)\in\Omega(u)\\ t\leqslant\sigma\leqslant t+T}} \varepsilon\{\lambda(t) f_x[\bar{X}(t)-\bar{x}(t)] + f_u[\bar{U}(t)-\bar{u}(t)]\} + \\ + \int_{t+\varepsilon}^{t+T} \{\lambda(\sigma) f_x[\bar{X}(\sigma)-\bar{x}(\sigma)] + f_u[\bar{U}(\sigma)-\bar{u}(\sigma)]\}\, d\sigma. \quad (3.204)$$

Mais la valeur minimale du deuxième terme est égale à $S[\bar{x}(t+\varepsilon), t+\varepsilon]$:

$$S(\bar{x}(t+\varepsilon), t+\varepsilon] = \\ = \min_{\substack{u(\sigma)\in\Omega(u)\\ t+\varepsilon\leqslant\sigma\leqslant t+T}} \int_{t+\varepsilon}^{t+T} \{\lambda(\sigma) f_x[\bar{X}(\sigma)-\bar{x}(\sigma)] + f_u[\bar{U}(\sigma)-\bar{u}(\sigma)]\}\, d\sigma. \quad (3.205)$$

Le deuxième terme de (3.204) peut être remplacé par $S[\bar{x}(t+\varepsilon), t+\varepsilon]$ sous la condition d'assurer la commande optimale dans l'intervalle $t+\varepsilon \leqslant \sigma \leqslant t+T$. Maintenant il faut trouver sur le petit intervalle $t \leqslant \sigma \leqslant t+\varepsilon$ une commande optimale $u(\sigma)$ telle qu'elle minimise le second membre de l'expression

$$S[\bar{x}(t), t] = \min_{u(t)\in\Omega(u)} [\varepsilon\{\lambda(t) f_x[\bar{X}(t)-\bar{x}(t)] + f_u[\bar{U}(t)-\bar{u}(t)]\} + \\ + S[\bar{x}(t+\varepsilon), t+\varepsilon]]. \quad (3.206)$$

Dans l'expression (3.206) il est facile de passer à la limite pour $\varepsilon \to 0$. Procédant de même que dans le chapitre II et développant $S[\bar{x}(t+\varepsilon), t+\varepsilon]$ en série de Taylor, on obtient

$$S[x_1(t+\varepsilon), \ldots, x_n(t+\varepsilon), t+\varepsilon] = S[\bar{x}(t), t] + \\ + \frac{\partial S[\bar{x}(t), t]}{\partial t}\varepsilon + \sum_{k=1}^{n} \frac{\partial S[\bar{x}(t), t]}{\partial x_k}[x_k(t+\varepsilon) - x_k(t)]. \quad (3.207)$$

Nous avons négligé ici les puissances supérieures de ε. Tout en négligeant toujours les puissances supérieures de ε, posons

$$x_k(t+\varepsilon) = x_k(t) + \varepsilon x_k'(t), \quad (3.208)$$

où $x_k'(t)$ est la dérivée de x_k par rapport au temps. Soit

$$S_t = \frac{\partial S}{\partial t}, \quad S_{x_k} = \frac{\partial S}{\partial x_k}; \quad (3.209)$$

on tire alors de (3.207)-(3.209)

$$S[\bar{x}(t+\varepsilon), t+\varepsilon] = S[\bar{x}(t), t] + \varepsilon S_t[\bar{x}(t), t] + \\ + \varepsilon \sum_{k=1}^{n} x_k'(t) S_{x_k}[\bar{x}(t), t]. \quad (3.210)$$

En portant (3.210) dans (3.205) on obtient

$$S[\bar{x}(t), t] = \min_{u(t)\in\Omega(u)} \{\varepsilon\{\lambda(t) f_x[\bar{X}(t) - \bar{x}(t)] + f_u[\bar{U}(t) - \bar{u}(t)]\} + \\ + S[\bar{x}(t), t] + \varepsilon S_t[\bar{x}(t), t] + \varepsilon \sum_{k=1}^{n} x_k'(t) S_{x_k}[\bar{x}(t), t]\}. \quad (3.211)$$

En éliminant $S[\bar{x}(t), t]$ des deux membres de l'égalité et en simplifiant par ε on aboutit à la condition nécessaire de la commande optimale sous la forme de l'égalité [cf. (2.199)]

$$\min_{u(t)\in\Omega(u)} \{\lambda(t) f_x[\bar{X}(t) - \bar{x}(t)] + f_u[\bar{U}(t) - \bar{u}(t)] + \\ + S_t[\bar{x}(t), t] + \sum_{k=1}^{n} x_k'(t) S_{x_k}[\bar{x}(t), t]\} = 0. \quad (3.212)$$

Des considérations physiques rendent clair que la fonction $\lambda(\sigma) \geqslant 0$ tandis que les fonctions f_k et f_u doivent être strictement convexes, par exemple $f_x(v) = f_u(v) = v^2$.

Examinons à titre d'illustration l'exemple très simple d'un objet caractérisé par une seule coordonnée x, alors que la réponse impulsionnelle $g(t, \tau)$ est de la forme

$$g(t, \tau) = g_1 \exp\left\{-\frac{t-\tau}{T_1}\right\}. \quad (3.213)$$

Ainsi l'objet représente un circuit apériodique dont le gain est $T_1 g_1$, la constante de temps T_1 et l'équation

$$x'(t) + \frac{1}{T_1} x(t) = g_1 u(t). \quad (3.214)$$

Soit ensuite $f_x(v) = f_u(v) = v^2$. La condition (3.212) s'écrit alors

$$\min_{u(t)\in\Omega(u)} \{\lambda(t)[X(t) - x(t)]^2 + [U(t) - u(t)]^2 + \\ + S_t[x(t), t] + x'(t) S_x[x(t), t]\} = 0. \quad (3.215)$$

Si dans (3.215) on substitue à la valeur $x'(t)$ sa valeur tirée de (3.214), on obtient l'égalité

$$\min_{u(t)\in\Omega(u)} \left\{\lambda(t)[X(t) - x(t)]^2 + [U(t) - u(t)]^2 + \\ + S_t[x(t), t] + \left[g_1 u(t) - \frac{1}{T_1} x(t)\right] S_x[x(t), t]\right\} = 0. \quad (3.216)$$

Supposons d'abord que les contraintes limitant la commande $u(t)$ soient inexistantes et qu'elle puisse varier dans les limites $-\infty < u(t) < \infty$. On peut alors minimiser l'expression (3.216) en la dérivant par rapport à u et en égalant la dérivée à zéro

$$-2\,[U(t) - u(t)] + g_1 S_x\,[x(t),\ t] = 0. \tag{3.217}$$

On en tire la commande optimale

$$u^*(t) = U(t) - \frac{g_1}{2} S_x\,[x(t),\ t]. \tag{3.218}$$

Ainsi pour rechercher la commande optimale $u^*(t)$ il faut connaître la fonction $S\,[x(t),\ t]$. En portant (3.218) dans (3.216), on obtient

$$\lambda(t)\,[X(t) - x(t)]^2 + S_t\,[x(t),\ t] + \\ + \left[g_1 U(t) - \frac{g_1^2}{2} S_x\,[x(t),\ t] - \frac{1}{T_1}\,x(t)\right] S_x\,[x(t),\ t] - \\ - \left\{\frac{g_1}{2} S_x\,[x(t),\ t]\right\}^2 = 0, \tag{3.219}$$

qui est une équation différentielle aux dérivées partielles. Il faut trouver la solution de l'équation (3.219) qui vérifie la condition aux limites (3.203).

Remarquons que la plus grande puissance que présente $x(t)$ dans l'équation (3.219) est égale à deux [le premier membre de l'équation comporte le terme $\lambda(t)\,x^2(t)$]. Ceci fait penser qu'en développant la fonction $S\,(x(t),\ t]$ en série par rapport à x il suffit également de se borner par la puissance deux. Par conséquent, posons

$$S\,[x(t),\ t] = K(t) + K_1(t)\,x(t) + K_{11}(t)\,[x(t)]^2, \tag{3.220}$$

où $K(t)$, $K_1(t)$ et $K_{11}(t)$ sont certaines fonctions du temps t qu'il faut définir par substitution de la solution hypothétique (3.220) dans (3.219). Calculons d'abord $S_t\,[x(t),\ t]$ et $S_x\,[x(t),\ t]$ en dérivant (3.220) par rapport à t et à x respectivement. On est ainsi amené à

$$S_t\,[x(t),\ t] = K'(t) + K_1'(t)\,x(t) + K_{11}'(t)\,[x(t)]^2. \tag{3.221}$$

L'apostrophe désigne ici la dérivée par rapport à t de la fonction respective. Ensuite

$$S_x\,[x(t),\ t] = K_1(t) + 2K_{11}(t)\,x(t). \tag{3.222}$$

Portons maintenant (3.220), (3.221) et (3.222) dans (3.219) et groupons les termes de puissances égales de $x(t)$:

$$\left\{K'(t)+g_1K_1(t)U(t)-\left[\frac{g_1}{2}K_1(t)\right]^2+\lambda(t)[X(t)]^2\right\}+$$
$$+x(t)\left\{K_1'(t)+2g_1K_{11}(t)U(t)-\frac{1}{T_1}K_1(t)-\right.$$
$$\left.-\frac{g_1^2}{2}K_1(t)K_{11}(t)-2\lambda(t)X(t)\right\}+$$
$$+[x(t)]^2\left\{K_{11}'(t)-\frac{2}{T_1}K_{11}(t)-[g_1K_{11}(t)]^2+\lambda(t)\right\}=0. \quad (3.223)$$

Etant donné que l'équation (3.223) doit être juste pour toutes les valeurs de $x(t)$, les coefficients affectés à $x(t)$ de puissances nulle, première et deuxième doivent être nuls. On est amené ainsi aux trois équations différentielles ordinaires pour les fonctions $K(t)$, $K_1(t)$ et $K_{11}(t)$:

$$K'(t)=\left[\frac{g_1}{2}K_1(t)\right]^2-g_1K_1(t)U(t)-\lambda(t)[X(t)]^2, \quad (3.224)$$

$$K_1'(t)=\left[\frac{1}{T_1}+\frac{g_1^2}{2}K_{11}(t)\right]K_1(t)-2g_1K_{11}(t)U(t)+2\lambda(t)X(t) \quad (3.225)$$

et

$$K_{11}'(t)=\frac{2}{T_1}K_{11}(t)+[g_1K_{11}(t)]^2-\lambda(t). \quad (3.226)$$

Pour résoudre ces équations différentielles il faut encore connaître les valeurs initiales ou les valeurs frontière des fonctions $K(t)$, $K_1(t)$ et $K_{11}(t)$. Si l'on compare les conditions (3.203) avec l'égalité (3.220) on voit que (3.203) n'est vérifiée pour n'importe quels $x(t+T)$ finaux que dans le cas où l'on admet que les coefficients $K_1(t+T)$ et $K_{11}(t+T)$ sont égaux à zéro :

$$K_1(t+T)=K_{11}(t+T)=0. \quad (3.227)$$

Mais alors en vertu de (3.203) on a

$$K(t+T)=0. \quad (3.228)$$

Il en résulte que les solutions des équations (3.224)-(3.226) doivent satisfaire aux trois conditions aux limites (3.227) et (3.228).

Lorsque les fonctions $K(t)$, $K_1(t)$ et $K_{11}(t)$ sont obtenues, la recherche de la loi de commande optimale $u^*(t)$ ne présente plus aucune difficulté. En portant dans (3.218) S_x tiré de (3.222) on arrive à

$$u^*(t)=\left[U(t)-\frac{g_1}{2}K_1(t)\right]-g_1K_{11}(t)x(t). \quad (3.229)$$

Ainsi $u^*(t)$ dépend du temps t non seulement explicitement par les fonctions $K_1(t)$, $K_{11}(t)$ et $U(t)$, mais encore implicitement par $x(t)$.

Pour résoudre les équations (3.224)-(3.226) avec les conditions aux limites (3.227) et (3.228) on peut faire appel à l'inversion du temps en considérant que $(t+T)$ est l'instant initial. Les conditions (3.227) et (3.228) deviennent alors « initiales » et on se retrouve devant un problème de Cauchy usuel avec des conditions « initiales » imposées. Il est impossible d'obtenir la solution de ces équations sous une forme finie car elles sont non linéaires. Pourtant on peut les résoudre d'une manière usuelle sur des calculateurs.

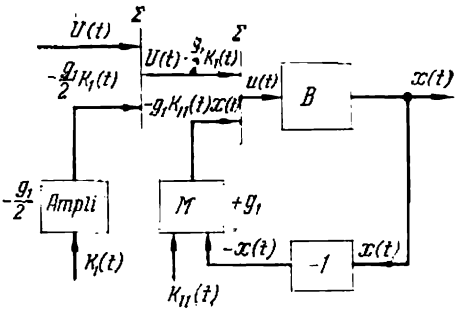


Fig. 3.23.

La figure 3.23 représente le schéma fonctionnel d'un organe de commande régi par l'équation (3.229). On a montré les fonctions $K_1(t)$ et $K_{11}(t)$ amenées à partir des calculateurs résolvant les équations (3.224)-(3.226). La fonction $K_1(t)$ en passant par l'amplificateur de gain $-\frac{g_1}{2}$ donne le terme $-\frac{g_1}{2}K_1(t)$ qui s'ajoute dans l'additionneur $\sum$ à $U(t)$ appliquée de l'extérieur pour donner le terme entre crochets du second membre de (3.229). Le deuxième terme de l'expression (3.229) est fourni par le multiplieur M. La grandeur de sortie de ce dernier est égale au produit des grandeurs d'entrée $K_{11}(t)$ et $-x(t)$ multiplié par le coefficient constant g_1. Le facteur $-x(t)$ s'obtient à la sortie de l'inverseur $\boxed{-1}$, $x(t)$ étant fourni à l'entrée de l'inverseur à partir de l'objet gouverné B.

Considérons maintenant le cas de $u(t)$ obéissant à la contrainte

$$L_m^-(t) \leqslant u(t) \leqslant L_m^+(t).$$

On ne peut plus tirer de (3.219) la valeur de u minimisant le premier membre de (3.216) puisque on risque de tomber sur une valeur de u^* qui dépasse les limites imposées. Evidemment, les cas possibles sont les suivants.

a) La solution $u_0(t)$ de l'équation (3.229), c'est-à-dire l'expression

$$u_0(t) = \left[U(t) - \frac{g_1}{2}K_1(t)\right] - g_1K_{11}(t)\,x(t), \qquad (3.230)$$

se trouve dans les limites imposées. Par conséquent les expressions précédentes demeurent inchangées.

b) La solution de (3.230) dépasse la limite supérieure de $u(t)$

$$u_0(t) > L_m^+(t). \tag{3.231}$$

Nous sommes dans le cas où la valeur de $u^*(t)$ minimisant le premier membre de (3.216) s'écrit

$$u^*(t) = L_m^+(t). \tag{3.232}$$

c) La solution de (3.230) est plus petite que la limite inférieure

$$u_0(t) < L_m^-(t). \tag{3.233}$$

Le premier membre de (3.216) atteint alors le minimum pour

$$u^*(t) = L_m^-(t). \tag{3.234}$$

La réunion de tous les cas fournit la loi de commande

$$u^*(t) = \begin{cases} L_m^+(t), & u_0(t) \geqslant L_m^+(t), \\ u_0(t), & L_m^-(t) \leqslant u_0(t) \leqslant L_m^+(t), \\ L_m^-(t), & u_0 \leqslant L_m^-(t). \end{cases} \tag{3.235}$$

Il importe pourtant de noter que dans ce cas les fonctions $K(t)$, $K_1(t)$ et $K_{11}(t)$ ne peuvent plus s'obtenir à partir des anciennes équations (3.224)-(3.226). En effet, l'équation (3.219), qui a servi à déterminer ces fonctions, n'est pas applicable pour les instants où $u_0(t)$ sort des limites admissibles. Si, par exemple, la condition (3.231) est vérifiée pour un instant quelconque, donc l'expression (3.232) l'est également. L'équation aux dérivées partielles pour S peut alors s'obtenir en remplaçant $u(t)$ dans le premier membre de (3.216) par l'expression $L_m^+(t)$. Dans ce cas on obtient l'équation

$$\lambda(t)[X(t)-x(t)]^2+[U(t)-L_m^+(t)]^2+ \\ +S_t[x(t), t]+\left[g_1L_m^+(t)-\frac{1}{T_1}x(t)\right]\cdot S_x[x(t), t]=0. \tag{3.236}$$

Portons dans cette équation les formules (3.221) et (3.222), il vient :

$$\{K'(t)+g_1L_m^+(t)K_1(t)+[U(t)-L_m^+(t)]^2+\lambda(t)[X(t)]^2\}+ \\ +x(t)\left\{K_1'(t)+2g_1L_m^+(t)K_{11}(t)-\frac{1}{T_1}K_1(t)-2\lambda(t)X(t)\right\}+ \\ +[x(t)]^2\left\{K_{11}'(t)+\frac{2}{T_1}K_{11}(t)+\lambda(t)\right\}=0. \tag{3.237}$$

Si les coefficients affectés aux puissances zéro, une et deux de $x(t)$ sont égaux à zéro, cette équation est vérifiée quel que soit $x(t)$. Ceci nous donne trois équations pour les fonctions $K(t)$, $K_1(t)$

et $K_{11}(t)$:

$$\left.\begin{aligned} K'(t) &= -g_1 L_m^+(t) K_1(t) - \{U(t) - L_m^+(t)\} - \lambda(t)[X(t)]^2, \\ K_1'(t) &= \frac{1}{T} K_1(t) - 2g_1 L_m^+(t) K_{11}(t) + 2\lambda(t) X(t), \\ K_{11}'(t) &= \frac{2}{T_1} K_{11}(t) - \lambda(t). \end{aligned}\right\} \quad (3.238)$$

Des équations analogues s'obtiennent en remplaçant $L_m^+(t)$ par $L_m^-(t)$ lorsque l'inégalité vérifiée est (3.233) et la commande optimale est fournie par la formule (3.234).

Ainsi les fonctions $K(t)$, $K_1(t)$ et $K_{11}(t)$ s'obtiennent à partir d'un système d'équations formé par réunion de tous les trois systèmes d'équations valables chacun dans l'un des trois cas possibles a), b) et c) énumérés ci-dessus. Au lieu de ces systèmes d'équations on peut certainement écrire un système général unique dont les coefficients varient suivant que $u_0(t)$ se trouve à l'intérieur de l'intervalle admissible ou le dépasse dans l'un ou l'autre sens. Ces équations s'intègrent pour les mêmes conditions aux limites (3.227) et (3.228). Pourtant pour réaliser maintenant l'intégration rapide de ces équations avec inversion du temps, il faut adjoindre un calculateur analogique sur lequel est intégrée, également à rebours, l'équation de l'objet (3.214) : de plus, la valeur courante de $x(t)$ sera désormais « finale ». L'intégration de cette équation est nécessaire du fait que d'après (3.230) les valeurs de $u_0(\sigma)$, $t < \sigma < t + T$, dépendent de $x(\sigma)$. C'est pourquoi la définition de la valeur de la commande optimale $u^*(t)$ à l'instant courant impose un calcul rapide préalable des fonctions $K(t)$, $K_1(t)$ et $K_{11}(t)$ réalisé sur un calculateur analogique qui « prévoit » également les valeurs futures de ces fonctions de même que de $u^*(\sigma)$ et de $x(\sigma)$. Lorsque les valeurs de ces fonctions sont trouvées, la commande est réalisée d'après la formule (3.235), alors que $u_0(t)$ est fournie par (3.230).

En général, quel que soit l'objet linéaire et le critère d'optimalité quadratique, même lorsque la commande $u(t)$ est bornée à un intervalle admissible, il existe une formule précise de commande optimale $u^*(t)$, composée d'un nombre fini de termes. D'autres critères d'optimalité conduisent aux lois de commande optimale dont l'approximation ne peut s'obtenir avec un nombre fini de termes que d'une manière grossière.

La méthode examinée ci-dessus peut être généralisée aux systèmes d'ordre plus élevé [3.34-3.36].

La programmation dynamique est un procédé très général de formulation et de préparation de la résolution approchée des problèmes les plus variés. Ainsi pour l'objet dont l'équation s'écrit

$$\frac{dx}{dt} = f(x, u), \quad (3.239)$$

et qui est soumis à la condition initiale

$$(x)_{t=0} = x^{(0)}, \tag{3.240}$$

on peut demander de trouver la commande optimale $u^* \in \Omega(u)$ telle qu'elle minimise le maximum de l'écart $x(t)$ par rapport à une certaine fonction connue $x^*(t)$ dans l'intervalle $0 \leqslant t \leqslant T$, où T est fixé. Dans le cas concerné le critère d'optimalité est de la forme

$$Q = \max_{0 \leqslant t \leqslant T} |x^*(t) - x(t)|. \tag{3.241}$$

Dans cet intervalle, $u(t)$ doit être telle qu'elle vérifie la condition

$$\min Q = \min_{\substack{u(t) \in \Omega(u) \\ 0 \leqslant t \leqslant T}} \max |x^*(t) - x(t)| = S(x^{(0)}). \tag{3.242}$$

La lettre S désigne la valeur minimale de Q définie par la condition initiale $x^{(0)}$. La commande $u(t)$ appartient à un certain domaine admissible $\Omega(u)$.

Pour résoudre le problème (cf. [2.9]) remplaçons l'équation différentielle (3.239) par l'équation aux différences finies approchée

$$x_{k+1} = x_k + f(x_k, u_k)\,\Delta, \tag{3.243}$$

où

$$x_k = x(k\Delta), \quad u_k = u(k\Delta), \quad \Delta = \frac{T}{N} \tag{3.244}$$

et N est le nombre total des intervalles élémentaires contenus dans l'intervalle de temps T. Au lieu de (3.241) on peut maintenant écrire

$$Q = \max_{k=0,\,1,\,\ldots,\,N} |x_k^* - x_k|, \tag{3.245}$$

alors que l'expression (3.242) devient

$$S(x^{(0)}) = \min Q = \min_{u_j \in \Omega(u),\ k=0,\,1,\,\ldots,\,N} \max_{j=0,\,1,\,\ldots,\,N-1} |x_k^* - x_k|. \tag{3.246}$$

Considérons d'abord un seul instant $t = N$. Pour cet instant, avec $x = x^{(0)}$, on a

$$S_N(x_N) = \min Q_N = |x_N^* - x_N|, \tag{3.247}$$

et la variation de la commande u n'influe en général pas sur la valeur de Q_N. Déplaçons-nous à rebours depuis l'instant $t = N\Delta$ vers $t = N - 1$. Examinons la fonction

$$Q_{N-1} = \max_{k=N-1,\,N} |x_k^* - x_k| = \max\{|x_{N-1}^* - x_{N-1}|,\ |x_N^* - x_N|\} =$$
$$= \max\{|x_{N-1}^* - x_{N-1}|,\ |x_N^* - x_{N-1} - f(x_{N-1}, u_{N-1})\,\Delta|\}. \tag{3.248}$$

Ici x_N est remplacé d'après (3.243). La notation $\max\{a, b\}$ signifie : la valeur maximale des deux grandeurs a et b. Avec la valeur donnée

de x_{N-1}, la grandeur Q_{N-1} dépend de u_{N-1}. Choisissons la valeur de $u^*_{N-1} \in \Omega(u)$ telle qu'elle minimise Q^*_{N-1}. On obtient alors la commande optimale u^*_{N-1}, fonction de x_{N-1}, x^*_{N-1}, x^*_N, ainsi que la valeur minimale de Q_{N-1} :

$$S_{N-1}(x_{N-1}) = \min_{u_{N-1} \in \Omega(u)} Q_{N-1}. \tag{3.249}$$

Dans ces conditions il se peut que la variation de u_{N-1} dans une certaine marge n'influe pas du tout sur (3.248), pour peu que la première des valeurs dans l'accolade soit supérieure à la deuxième. On en déduit que la commande optimale est définie d'une façon non univoque. Convenons de définir toujours la valeur de u minimisant le deuxième terme dans l'accolade (3.248).

Passons maintenant à l'instant $t = (N-2)\,\Delta$. Dans l'intervalle de $t = (N-2)\,\Delta$ à $t = N\Delta$ on a

$$Q_{N-2} = \max_{k=N-2,\,N-1,\,N} |x^*_k - x_k| = \max\{|x^*_{N-2} - x_{N-2}|,$$
$$\max\{|x^*_{N-1} - x_{N-1}|,\ |x^*_N - x_N|\}\}. \tag{3.250}$$

La commande u_{N-1} n'influe que sur le deuxième terme dans l'accolade extérieure de cette expression. Si on la choisit toujours d'après la loi (3.249) établie plus haut, il vient

$$S_{N-2}(x_{N-2}) = \min_{u_{N-1},\,u_{N-2} \in \Omega(u)} Q_{N-2} = \min_{u_{N-2} \in \Omega(u)} \max\{|x^*_{N-2} - x_{N-2}|,$$
$$S_{N-1}(x_{N-1})\} =$$
$$= \min_{u_{N-2} \in \Omega(u)} \max\{|x^*_{N-2} - x_{N-2}|,\ S_{N-1}(x_{N-2} + f(x_{N-2}, u_{N-2})\,\Delta)\} =$$
$$= \max\{|x^*_{N-2} - x_{N-2}|,\ \min_{u_{N-2} \in \Omega(u)} S_{N-1}(x_{N-2} + f(x_{N-2}, u_{N-2})\,\Delta)\}. \tag{3.251}$$

Après avoir minimisé le second membre de (3.251), on tombe sur la commande optimale u^*_{N-2}, etc. En reprenant ce raisonnement on calcule par récurrence les fonctions $S_{N-k}(x_{N-k})$ et en même temps la commande optimale u^*_{N-k}

$$S_{N-k}(x_{N-k}) = \max\{|x^*_{N-k} - x_{N-k}|,$$
$$\min_{u_{N-k} \in \Omega(u)} S_{N-k+1}(x_{N-k} + f(x_{N-k}, u_{N-k}) \cdot \Delta\}. \tag{3.252}$$

Ici $k = 1, 2, \ldots, N$. La fonction $S_N(x_N)$ étant connue [cf. (3.247)], on peut en principe trouver toutes les S_{N-k} fonctions, y compris $S_0(x_0) = S_0(x^{(0)})$. La commande correspondante u_0 est précisément celle que nous recherchons.

La réalisation de cette procédure n'est possible que sous une forme numérique et n'impose pas de calculs excessivement volumineux.

La programmation dynamique permet de définir la commande optimale lorsque le temps T du processus est non fixé, y compris le cas de la commande en temps minimal. Soit, par exemple, l'équation de l'objet de la forme

$$a_1 x + \sum_{i=1}^{n} a_{i+1} \frac{d^i x}{dt^i} = u(t). \tag{3.253}$$

Posons

$$x = x_1, \quad \frac{d^i x}{dt^i} = x_{i+1} \quad (i = 1, 2, \ldots, n-1). \tag{3.254}$$

L'équation (3.253) se décompose en n équations de premier ordre:

$$\left.\begin{aligned} &\frac{dx_i}{dt} = x_{i+1} \quad (i = 1, 2, \ldots, n-1), \\ &\frac{dx_n}{dt} = -\sum_{i=1}^{n} \frac{a_i}{a_{n+1}} x_i + \frac{u}{a_{n+1}} = -\sum_{i=1}^{n} b_i x_i + v, \end{aligned}\right\} \tag{3.255}$$

où

$$b_i = \frac{a_i}{a_{n+1}}, \quad \frac{u}{a_{n+1}} = v. \tag{3.256}$$

Soit v borné en module

$$|v| \leqslant V. \tag{3.257}$$

Recherchons la commande optimale $v(t)$ telle qu'elle minimise le temps T du déplacement du point représentatif depuis la position initiale $x_1^{(0)}, \ldots, x_n^{(0)}$, donnée par le vecteur $\bar{x}^{(0)}$, jusqu'à l'origine des coordonnées dans l'espace de phase.

Remplaçons l'équation (3.255) par des équations aux différences finies, en posant $T = N\Delta t$, $x_i^{(k)} = x_i(k\,\Delta t)$, $v^{(k)} = v(k\,\Delta t)$:

$$\left.\begin{aligned} &x_i^{(k+1)} = x_i^{(k)} + \Delta t \cdot x_{i+1}^{(k)} \quad (i = 1, 2, \ldots, n-1), \\ &x_n^{(k+1)} = x_n^{(k)} - \Delta t \sum_{i=1}^{n} b_i x_i^{(k)} + \Delta t \cdot v^{(k)}. \end{aligned}\right\} \tag{3.258}$$

La relation définissant le temps minimal nécessaire au point représentatif pour entrer à l'intérieur d'une sphère de petit diamètre δ, lorsque $k = N$, s'écrit:

$$\sum_{i=1}^{n} [x_i^{(N)}]^2 \leqslant \delta^2. \tag{3.259}$$

Le temps minimal T ne dépend que des conditions initiales, c'est-à-dire du vecteur $\bar{x}^{(0)}$. Il s'ensuit que $T = T(\bar{x}^{(0)})$. Le déplacement d'un pas à partir de la position initiale se fait dans le temps

Δt, après quoi le vecteur $\bar{x}^{(0)}$ est remplacé par le vecteur $\bar{x}^{(1)}$ dépendant de $v^{(0)}$. Le temps minimal nécessaire pour atteindre à partir du point $(\bar{x}^{(1)})$ la petite sphère est égal à $T\ (x^{(1)})$, alors que le temps total est $\Delta t + T\ (\bar{x}^{(1)})$. Il est clair que

$$T\ (\bar{x}^{(0)} = \min_{|v^{(0)}|\leqslant V} [\Delta t + T\ (\bar{x}^{(1)})] = \Delta t + \min_{|v^{(0)}|\leqslant V} [T\ (\bar{x}^{(1)})]. \quad (3.260)$$

Si l'expression de $\bar{x}^{(1)}$ est développée d'après (3.258), il vient

$$T\ (x_1^{(0)}, \ldots, x_n^{(0)}) = \Delta t + \min_{|v^{(0)}|\leqslant V} T\left(x_1^{(0)} + \Delta t \cdot x_2^{(0)}, \ldots, x_{n-1}^{(0)} + \right.$$

$$\left. + \Delta t \cdot x_n^{(0)}, x_n^{(0)} - \Delta t \sum_{i=1}^{n} b_i x_i^{(0)} + \Delta t \cdot v^{(0)}\right). \quad (3.261)$$

Supposons que T est une fonction dérivable des variables $x_i^{(0)}$. On a

$$T\ (x_1^{(0)}, \ldots, x_n^{(0)}) = \Delta t + \min_{|v^{(0)}|\leqslant V} \left\{T\ (x_1^{(0)}, \ldots, x_n^{(0)}) + \sum_{i=1}^{n-1} \frac{\partial T}{\partial x_i^{(0)}} \Delta t \cdot x_{i+1}^{(0)} + \right.$$

$$\left. + \frac{\partial T}{\partial x_n^{(0)}} \left(-\Delta t \sum_{i=1}^{n} b_i x_i^{(0)} + \Delta t v^{(0)}\right)\right\} + o\ (\Delta t). \quad (3.262)$$

Il n'y a que le dernier terme dans l'accolade qui dépend de $v^{(0)}$. Par conséquent, ce n'est que sur ce dernier que porte la minimisation par rapport à $v^{(0)}$. Supprimons dans les deux membres de l'égalité (3.262) le terme $T\ (x_1^{(0)}, \ldots, x_n^{(0)})$, puis, après avoir simplifié par Δt et faisant tendre Δt vers zéro, nous aboutissons à l'expression

$$0 = 1 + \sum_{i=1}^{n-1} \frac{\partial T}{\partial x_i^{(0)}} x_{i+1}^{(0)} + \min_{|v^{(0)}|\leqslant V} \left\{\frac{\partial T}{\partial x_n^{(0)}} \left(-\sum_{i=1}^{n} b_i x_i^{(0)} + v^{(0)}\right)\right\}. \quad (3.263)$$

Avec $\Delta t \to 0$, il convient également de faire tendre vers zéro la valeur de δ figurant dans (3.259). En remplaçant les $x_i^{(0)}$ par les valeurs courantes x_i, que nous pouvons toujours considérer comme initiales, nous tombons sur l'équation différentielle aux dérivées partielles définissant $T\ (x_1, x_2, \ldots, x_n)$:

$$0 = 1 + \sum_{i=1}^{n-1} \frac{\partial T}{\partial x_i} x_{i+1} - \frac{\partial T}{\partial x_n} \sum_{i=1}^{n} b_i x_i + \min_{|v|\leqslant V} \left\{\frac{\partial T}{\partial x_n} \cdot v\right\}. \quad (3.264)$$

En comparant (3.257) et (3.264) on constate sans peine que le minimum du dernier terme du second membre de (3.264) s'obtient lorsqu'est vérifiée la condition

$$v^* = -V \operatorname{sign} \frac{\partial T}{\partial x_n}. \quad (3.265)$$

C'est précisément la loi de la commande optimale. On en déduit qu'il faut toujours choisir les valeurs de $v(t)$ aux frontières du domaine admissible

$$v^* = \pm V. \tag{3.266}$$

Nous l'avons déjà dit d'ailleurs au § 1 du présent chapitre.

En portant (3.265) dans (3.264) on a

$$0 = 1 + \sum_{i=1}^{n-1} \frac{\partial T}{\partial x_i} x_{i+1} - \frac{\partial T}{\partial x_n} \sum_{i=1}^{n} b_i x_i - V \left| \frac{\partial T}{\partial x_n} \right|. \tag{3.267}$$

La solution de l'équation (3.267) doit vérifier la condition aux limites

$$T(0, 0, \ldots, 0) = 0, \tag{3.268}$$

puisqu'avec la valeur initiale $\bar{x}^{(0)} = 0$, le temps nécessaire pour rejoindre l'origine des coordonnées est évidemment égal à zéro.

Pour obtenir une commande optimale v^* en fonction de x, il faut, comme on le voit de (3.265), trouver la fonction $T(\bar{x})$, c'est-à-dire résoudre l'équation aux dérivées partielles (3.267).

La solution explicite ne peut s'obtenir que pour les cas les plus simples. Considérons à titre d'exemple les équations de l'objet

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = u \tag{3.269}$$

avec la contrainte

$$|u| \leqslant 1. \tag{3.270}$$

Il faut passer en un temps minimal du point $(x_1^{(0)}, x_2^{(0)})$ à l'origine des coordonnées.

Si on remplace v par u, l'équation (3.264) se met sous la forme

$$0 = 1 + \frac{\partial T}{\partial x_1} x_2 + \min_{|u| \leqslant 1} \left\{ \frac{\partial T}{\partial x_2} u \right\}, \tag{3.271}$$

et la loi de la commande optimale s'écrit

$$u^* = -\operatorname{sign} \frac{\partial T}{\partial x_2}. \tag{3.272}$$

D'après (3.266) et (3.270), la commande optimale ne prend que les valeurs ± 1. Le plan de phase (x_1, x_2) doit alors être partagé en deux domaines (fig. 3.13): L_2' dans lequel $u = -1$ et L_2'' où $u = +1$. Pour le premier, L_2', (3.271) s'écrit

$$0 = 1 + \frac{\partial T}{\partial x_1} x_2 - \frac{\partial T}{\partial x_2} \tag{3.273}$$

et pour L_2'' on a

$$0 = 1 + \frac{\partial T}{\partial x_1} x_2 + \frac{\partial T}{\partial x_2}. \tag{3.274}$$

Pour résoudre ces équations il faut connaître les conditions aux limites sur l'une quelconque des lignes. On peut intégrer, par exemple, l'équation (3.269) considérant que $u(t)$ varie suivant une courbe « rectangulaire » $u = \pm 1$. On aboutit alors à $T(x_1)$ pour $x_2 = 0$.

On peut montrer (cf. [3.37]) que la solution de l'équation (3.273) est de la forme

$$T = 2\sqrt{\frac{1}{2}x_2^2 + x_1} + x_2, \tag{3.275}$$

et celle de l'équation (3.274)

$$T = 2\sqrt{\frac{1}{2}x_2^2 - x_1} - x_2. \tag{3.276}$$

En portant (3.272) dans (3.271) on arrive à

$$0 = 1 + \frac{\partial T}{\partial x_1}x_2 - \left|\frac{\partial T}{\partial x_2}\right|. \tag{3.277}$$

Les formules (3.275) et (3.276) prises à part ne constituent pas encore la solution de l'équation (3.277) Cette solution a la forme

$$T = \begin{cases} 2\sqrt{\frac{1}{2}x_2^2 + x_1} + x_2, & \bar{x} \in L_2', \\ 2\sqrt{\frac{1}{2}x_2^2 - x_1} - x_2, & \bar{x} \in L_2''. \end{cases} \tag{3.278}$$

Cherchons le domaine L_2'' dans lequel est vérifiée la formule (3.276). A cet effet, portons dans la formule (3.277) les dérivées partielles du second membre de (3.276), c'est-à-dire

$$\frac{\partial T}{\partial x_1} = \frac{-1}{\sqrt{\frac{1}{2}x_2^2 - x_1}}, \quad \frac{\partial T}{\partial x_2} = \frac{x_2}{\sqrt{\frac{1}{2}x_2^2 - x_1}} - 1. \tag{3.279}$$

Il vient

$$\left(1 - \frac{x_2}{\sqrt{\frac{1}{2}x_2^2 - x_1}}\right) - \left|\frac{x_2}{\sqrt{\frac{1}{2}x_2^2 - x_1}} - 1\right| = 0. \tag{3.280}$$

Cette égalité est vérifiée pour la condition

$$1 - \frac{x_2}{\sqrt{\frac{1}{2}x_2^2 - x_1}} \geqslant 0 \tag{3.281}$$

ou

$$\sqrt{\frac{1}{2}x_2^2 - x_1} \geqslant x_2. \tag{3.282}$$

Cette dernière condition est vérifiée avec

$$x_2 > 0, \quad x_1 + \frac{x_2^2}{2} \leqslant 0, \tag{3.283}$$

ou avec

$$x_2 < 0, \quad \frac{1}{2} x_2^2 - x_1 \geqslant 0, \tag{3.284}$$

puisque la radicande ne peut pas être négative. Les conditions (3.283) et (3.284) entraînent que la frontière du domaine L_2'' est donnée par les équations

$$\left.\begin{aligned} x_2 > 0, \quad & x_1 = -\frac{x_2^2}{2}, \\ x_2 < 0, \quad & x_1 = \frac{x_2^2}{2}, \end{aligned}\right\} \tag{3.285}$$

ou, sous une forme plus condensée,

$$x_1 = -\operatorname{sign} x_2 \frac{x_2^2}{2}. \tag{3.286}$$

La même frontière s'obtient pour L_2' si l'on utilise la formule (3.275).

La courbe de commutation étant donnée par (3.286), la confrontation de (3.283), (3.284) et (3.286) révèle que la loi optimale de la commutation s'écrit

$$u^* = -\operatorname{sign}\left(x_1 + \operatorname{sing} x_2 \frac{x_2^2}{2}\right). \tag{3.287}$$

Il est vrai que pour l'exemple considéré la méthode du plan de phase est bien plus simple que celle de programmation dynamique. Pourtant, cette dernière est bien plus générale. Nous montrons dans les chapitres qui suivent qu'elle est également applicable à l'étude des problèmes statistiques.

La programmation dynamique peut s'employer non seulement dans le cas des systèmes échantillonnés dans le temps (systèmes continus à états discrets), mais aussi pour les systèmes purement discrets, échantillonnés dans le temps et en amplitude [3.38]. Dans les systèmes de ce type les valeurs des grandeurs ne peuvent pas être quelconques ; elles doivent appartenir à un ensemble fini des niveaux admissibles.

On admet habituellement qu' à l'i-ième instant le système peut se trouver en état s_i. Il existe un nombre fini d'états possibles du système, désignés par $q_1, \ldots, q_m$ ou simplement par $1, \ldots, m$. L'état s_i peut être l'un des nombres $q_1, \ldots, q_m$, ce qui peut se noter, comme nous venons de le dire, par l'un des nombres $1, \ldots, m$. L'action extérieure u_i peut aussi prendre l'une des valeurs possibles.

Posons, pour plus de simplicité, qu'il n'existe que deux valeurs possibles de u_i, notamment 1 et 0. La loi de variation des états est donnée par la relation

$$s_{i+1} = f(s_i, u_i). \tag{3.288}$$

On admet souvent que la grandeur de sortie x_i est une fonction discrète de s_i et de u_i

$$x_i = \psi(s_i, u_i). \tag{3.289}$$

Pour simplifier, la grandeur x_i est identifiée dans ce qui suit à s_i, c'est-à-dire la relation (3.289) devient

$$x_i = s_i.$$

Le système discret est alors caractérisé par une seule fonction de deux variables x_i et u_i:

$$x_{i+1} = f(x_i, u_i), \tag{3.290}$$

que nous obtenons en remplaçant $s_i = x_i$ dans l'expression (3.288).

La relation (3.290) est souvent représentée sous la forme d'une table ou d'un graphe (cf., par exemple, [3.25], chapitre XV). Soit, par exemple, le système susceptible de prendre quatre états matérialisés par des cercles sur la figure 3.24 (les numéros des états sont indiqués par les chiffres placés à l'intérieur du cercle). Le passage d'un état à un autre est marqué sur le graphe par une flèche, à côté de laquelle on a figuré l'une des valeurs possibles de $u_i = 0, 1$ associée au passage considéré. Par exemple, la figure 3.24 montre que lorsque $s_i = 2$ et $u_i = 1$, le système passe à l'état $s_{i+1} = 3$; mais si $s_i = 2$ et $u_i = 0$, la flèche conduit de $s_i = 2$ à $s_{i+1} = 1$.

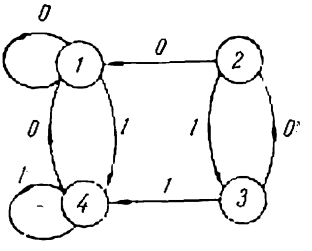


Fig. 3.24.

Il est possible que l'état du système ne change pas. Par exemple, si $s_i = 4$ et $u_i = 1$, l'état reste le même, ce qui est figuré par une flèche en boucle. Il en est de même pour $s_i = 1$ et $u_i = 0$.

u_i \ s_i	1	2	3	4
0	1	1	2	1
1	4	3	4	4

Cette relation peut être traduite par une table à deux entrées u_i et s_i en portant dans chaque case de cette table la valeur correspondante de s_{i+1}. La table qui correspond au graphe de la figure 3.24 est donnée ci-contre.

Par exemple, pour $s_i = 3$ et $u_i = 0$, on a $s_{i+1} = 2$.

Soit à l'instant initial $i = 0$ le système se trouve en un certain état initial s_0. Supposons que le processus soit considéré aux instants

discrets $i = 0, 1, \ldots, N$ où N est fixé. Il faut rechercher une suite des commandes $u_0, u_1, \ldots, u_N$ telle qu'elle maximise une certaine fonction donnée $\varphi(s_N)$ de l'état final s_N. C'est un problème typique de la théorie des systèmes optimaux, où le critère d'optimalité est

$$Q = \varphi(s_N). \tag{3.291}$$

Il est clair que la fonction $\varphi(s_N)$ doit être donnée sous la forme d'une table. Soit, par exemple, la fonction $\varphi(s_N)$ caractérisée par la table ci-contre.

s_N	1	2	3	4
$\varphi(s_N)$	-1	3	-2	2

Dans les systèmes discrets, la méthode de programmation dynamique se présente sous sa forme la plus pure. Il est commode de l'illustrer par un graphe (fig. 3.25).

Figurons d'abord, l'un sous l'autre, les cercles qui matérialisent quatre états possibles s_N à l'instant N. Ces cercles forment la colonne droite de la figure 3.25. La colonne du milieu représente quatre états possibles s_{N-1} à l'instant $N-1$; enfin, la première colonne à gauche indique les états s_{N-2} à l'instant $N-2$, etc. La figure 3.25 ne compte que trois colonnes; dans le cas général les colonnes sont au nombre de $N+1$ correspondant aux états $s_N, s_{N-1}, \ldots, s_0$.

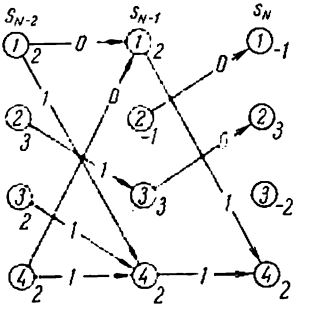


Fig. 3.25.

Commençons par la première colonne de droite. Près des états 1, 2, 3, 4 on a écrit les valeurs correspondantes de $\varphi(1)$, $\varphi(2)$, $\varphi(3)$, $\varphi(4)$ données par la table de $\varphi(s_N)$. Déplaçons-nous de s_N vers s_{N-1}. Soit, par exemple, l'état du système $s_{N-1} = 1$. Utilisant la table ou le graphe de la figure 3.24 cherchons une valeur de u_{N-1} telle qu'en partant de $s_{N-1} = 1$ on puisse obtenir la plus grande valeur possible de s_N. Avec $u_{N-1} = 1$ on a $s_N = 4$ à quoi correspond $\varphi(s_N) = 2$. Or, si l'on pose $u_{N-1} = 0$, il vient $s_N = 1$ et $\varphi(s_N) = -1$. On en déduit que la valeur optimale de u_{N-1} est 1; dans ce cas s'effectue la transition de $s_{N-1} = 1$ à $s_N = 4$. Cette transition assurée par une commande optimale est figurée par la ligne allant de $s_{N-1} = 1$ à $s_N = 4$. Le chiffre 1 marqué sur la flèche correspond à la commande optimale $u^*_{N-1} = 1$.

Traçons de la même façon les autres flèches qui relient les états s_{N-1} et s_N associés aux commandes optimales pour différentes valeurs « initiales » de S_{N-1} égales à 1; 2; 3; 4. Ces flèches portant les valeurs des commandes optimales relient chacun des états possibles s_{N-1} à l'un des états s_N. On assure ainsi la possibilité de maximiser le critère d'optimalité $\varphi(s_N)$ pour chacune des valeurs « initiales »

de s_{N-1}. Si par exemple, $s_{N-1} = 4$, la commande optimale $u^*_{N-1} = 1$ permet d'effectuer la transition en même position $s_N = 4$, après quoi la fonction $\varphi(s_N)$ devient égale à 2. Cette valeur du critère, maximale pour la condition « initiale » considérée $s_{N-1} = 4$, est notée près de l'état s_{N-1} par le chiffre 2. Les chiffres correspondants doivent être écrits près de chacune des valeurs de $s_{N-1} = 1, 2, 3, 4$.

Déplaçons-nous ensuite à rebours de s_{N-1} à s_{N-2} pour obtenir sur le même dessin les flèches reliant chacun des états possibles $s_{N-2} = 1, 2, 3, 4$ avec l'un des états s_{N-1} dans le cas d'une commande optimale u^*_{N-2}. Près des états « initiaux » s_{N-2} il faut porter les valeurs maximales possibles de $\varphi(s_N)$. En reprenant cette procédure pour s_{N-3}, s_{N-4}, etc., on arrive enfin à s_0, ayant construit un réseau de stratégies optimales correspondant à toutes les conditions initiales possibles.

Par exemple, avec $N = 2$, $s_{N-2} = s_0$. La figure 3.25 montre d'abord que dans ce cas pour la position initiale $s_0 = 3$ il faut choisir d'abord la commande optimale $u_{N-2} = u_0 = 1$ qui conduit à $s_{N-1} = s_1 = 3$, puis $u_{N-1} = u_1 = 0$, ce qui amène le système en $s_N = s_2 = 2$ et donne au critère $\varphi(s_N)$ la valeur maximale possible $Q = 3$.

La stratégie optimale n'est pas nécessairement univoque. Ainsi, en partant de $s_{N-2} = 1$ on peut emprunter l'un des deux chemins suivant que u_{N-2} est égal à 0 ou à 1. Si dans le premier de ces cas on pose $u_{N-1} = 1$ et dans le deuxième également $u_{N-1} = 1$, on aboutit à l'état $s_N = 4$ avec la valeur maximale possible de $\varphi(s_N) = 2$. Le même cas de non-univocité a lieu quand $s_{N-2} = 4$. Il existe donc ici deux stratégies optimales à valeurs égales du critère Q

$$u_{N-2} = 1, \quad u_{N-1} = 1 \quad \text{et} \quad u_{N-2} = 0, \quad u_{N-1} = 1.$$

Ces procédés peuvent être généralisés aux cas plus complexes, lorsque, par exemple, les suites des commandes $u_0, \ldots, u_N$ ne peuvent pas être quelconques, mais sont soumises à des conditions supplémentaires [3.38].

§ 4. Application du principe du maximum

Le principe du maximum permet souvent d'établir les traits caractéristiques des processus optimaux se déroulant dans des systèmes continus. Considérons à titre d'exemple le problème de minimisation de la durée du processus transitoire dans le cas d'un objet décrit par les équations

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n) + b_i u \quad (i = 1, 2, \ldots, n) \qquad (3.292)$$

ou, sous une forme vectorielle,

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}) + \bar{b} \cdot u, \tag{3.293}$$

u étant ici l'unique commande, et $\bar{x}$, $\bar{f}$ et $\bar{b}$ des vecteurs; de plus

$$\bar{b} = (b_1, b_2, \ldots, b_n). \tag{3.294}$$

Les fonctions f_i sont considérées dérivables par rapport à leurs arguments.

Soit la contrainte

$$|u| \leqslant 1. \tag{3.295}$$

Composons les équations adjointes

$$\frac{d\psi_i}{dt} = -\sum_{\alpha=1}^{n} \frac{\partial f_\alpha}{\partial x_i} \psi_\alpha. \tag{3.296}$$

Formons ensuite l'hamiltonien

$$H = \left\langle \frac{d\bar{x}}{dt}, \bar{\psi} \right\rangle = \langle \bar{f}(\bar{x}) + \bar{b}u, \bar{\psi} \rangle = \langle \bar{f}(\bar{x}), \bar{\psi} \rangle + u \langle \bar{b}, \bar{\psi} \rangle. \tag{3.297}$$

Comme seul le deuxième terme de cette expression dépend de u, il est évident *) que le maximum de H s'obtient compte tenu de (3.295) avec

$$u = \operatorname{sign} \langle \bar{b}, \bar{\psi} \rangle = \operatorname{sign} \sum_{i=1}^{n} b_i \psi_i(t). \tag{3.298}$$

Pour expliciter la loi de commande il faut rechercher les ψ_i pour chaque point $\bar{x}$ de l'espace de phase. Pourtant la formule (3.298) permet déjà de tirer la conclusion importante que pour la classe de problèmes considérée la commande doit reposer sur la frontière du domaine des valeurs possibles (3.295). Nous montrons dans ce qui suit que pour d'autres classes de problèmes cette condition peut ne pas se vérifier.

Dans le cas particulier important d'un objet linéaire les équations (3.292) deviennent

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij} x_j + b_i u \quad (i = 1, 2, \ldots, n). \tag{3.299}$$

Supposons que l'équation caractéristique de la trajectoire d'un mouvement libre de l'objet (pour $u = 0$) n'a que des racines réelles

*) On suppose ici que le cas $\langle b, \psi(t) \rangle = 0$ n'a pas lieu et, par conséquent, les commandes dites singulières sont inexistantes (cf. [2.17, 2.21]).

négatives. Dans ces conditions, les équations adjointes (3.296) s'écrivent

$$\frac{d\psi_i}{dt} = -\sum_{\alpha=1}^{n} a_{\alpha i}\psi_\alpha. \tag{3.300}$$

Ces équations sont adjointes par rapport aux équations (3.299) de l'objet. On sait que lorsque les racines de l'équation principale sont réelles et négatives, celles de l'équation caractéristique du système adjoint (3.300) le sont aussi. Par conséquent, les solutions $\psi_i(t)$ sont de la forme

$$\psi_i(t) = \sum_{j=1}^{n} C_{ij}e^{p_j t}, \tag{3.301}$$

où p_j sont réelles et négatives *), alors que les constantes C_{ij} sont définies en résolvant le problème aux limites. On conclut que le second membre de (3.298) est aussi une somme de n exponentielles

$$u = \operatorname{sign} \sum_{i=1}^{n} b_i \sum_{j=1}^{n} C_{ij}e^{p_j t} = \operatorname{sign} \sum_{j=1}^{n} \left(\sum_{i=1}^{n} b_j C_{ij}\right) e^{p_j t} = \operatorname{sign} \sum_{j=1}^{n} D_j e^{p_j t}, \tag{3.302}$$

où

$$D_j = \sum_{i=1}^{n} b_i C_{ij}. \tag{3.303}$$

Toutefois l'expression $\sum_{j=1}^{n} D_j e^{p_j t}$ passe par zéro pas plus de $(n-1)$ fois et ne compte par conséquent pas plus de n intervalles de signe constant. Il s'ensuit que dans le cas général la commande u compte aussi n intervalles de signe constant, ce qui achève la démonstration du théorème des n intervalles.

La commande ne repose pourtant pas toujours sur la frontière du domaine admissible. Considérons, par exemple, un objet défini par l'équation (3.299) avec les contraintes (3.295) dans le cas où le critère d'optimalité s'écrit (cf. [3.39]) :

$$Q = \int_0^\infty \left(\sum_{j=1}^{n} a_j x_j^2 + cu^2\right) dt, \tag{3.304}$$

où tous les a_i et c sont considérés positifs.

---

*) On peut considérer, sans affecter la généralité, que les racines p_j sont différentes.

Posons

$$\left.\begin{aligned} f_0(x_1, \ldots, x_n, u) &= \sum_{j=1}^{n} a_j x_j^2 + cu^2, \\ f_1(x_1, \ldots, x_n, u) &= \sum_{j=1}^{n} a_{ij} x_j + b_i u \\ (i = 1, 2, \ldots, n). & \end{aligned}\right\} \tag{3.305}$$

Ajoutons aux coordonnées $x_1, \ldots, x_n$ de l'objet encore une, x_0, donnée par l'équation

$$\frac{dx_0}{dt} = f_0(x_1, \ldots, x_n, u), \quad x_0(0) = 0. \tag{3.306}$$

D'après (3.304), la valeur de Q devient alors égale à la limite de $x_0(t)$ pour $t \to \infty$. Composons le système des équations adjointes

$$\frac{d\psi_0}{dt} = 0, \quad \frac{d\psi_i}{dt} = -\sum_{\alpha=0}^{n} \frac{\partial f_\alpha}{\partial x_i} \psi_\alpha = -2a_i \psi_0 x_i - \sum_{j=1}^{n} a_{ji} \psi_j$$
$$(i = 1, 2, \ldots, n). \tag{3.307}$$

Maintenant on peut écrire l'expression pour l'hamiltonien

$$H = \sum_{\alpha=0}^{n} \psi_\alpha \frac{dx_\alpha}{dt} = \sum_{\alpha=0}^{n} \psi_\alpha f_\alpha(x_1, \ldots, x_n, u) =$$
$$= \psi_0 \left(\sum_{i=1}^{n} a_i x_i^2 + cu^2\right) + \sum_{i=1}^{n} \psi_i \left(\sum_{j=1}^{n} a_{ij} x_j + b_i u\right). \tag{3.308}$$

Comme la grandeur u figure dans l'expression

$$H_1 = c\psi_0 u^2 + u \sum_{i=1}^{n} \psi_i b_i, \tag{3.309}$$

la condition du maximum de H coïncide donc avec la condition

$$\max_{|u| \leqslant 1} H_1 = \max_{|u| \leqslant 1} \left[c\psi_0 u^2 + u \sum_{i=1}^{n} \psi_i b_i\right] =$$
$$= \max_{|u| \leqslant 1} \left\{c\psi_0 \left[u + \frac{1}{2c\psi_0} \sum_{i=1}^{n} b_i \psi_i\right]^2 - \frac{1}{4c\psi_0} \left[\sum_{i=1}^{n} b_i \psi_i\right]^2\right\}. \tag{3.310}$$

En vertu de la première des conditions (3.307), la grandeur ψ_0 est constante. Puisque sa valeur peut être choisie parmi les nombres négatifs quelconques (cf. chapitre II), posons

$$\psi_0 = -1. \tag{3.311}$$

En portant cette expression dans (3.310) on découvre aisément que l'expression dans l'accolade est maximale lorsque, si c'est possible, le premier terme, négatif, devient nul ou est minimal en valeur absolue. On voit sans peine que la grandeur

$$\left[u-\frac{1}{2c}\sum_{i=1}^{n} b_i\psi_i\right]^2 \tag{3.312}$$

est égale à zéro ou est minimale en valeur absolue si, pour $|u| \leqslant 1$, on choisit u comme suit :

$$u=\begin{cases} \frac{1}{2c}\sum\limits_{i=1}^{n} b_i\psi_i, & \text{si } \left|\frac{1}{2c}\sum\limits_{i=1}^{n} b_i\psi_i\right| \leqslant 1, \\ 1, & \text{si } \frac{1}{2c}\sum\limits_{i=1}^{n} b_i\psi_i \geqslant 1, \\ -1, & \text{si } \frac{1}{2c}\sum\limits_{i=1}^{n} b_i\psi_i \leqslant 1. \end{cases} \tag{3.313}$$

Les $\psi_i(t)$ deviennent connues si l'on résout les équations adjointes (3.307). Mais pour ce faire il faut rechercher d'abord les valeurs initiales $\psi_i(0)$.

Supposons d'abord que $u(t)$ n'atteint pas ses valeurs frontières. En remplaçant dans les équations (3.299) et (3.307) $u(t)$ par la première expression de (3.313), on a

$$\left.\begin{aligned} \frac{dx_i}{dt} &= \sum_{j=1}^{n} a_{ij}x_j + \frac{b_i}{2c}\sum_{j=1}^{n} b_j\psi_i, \\ \frac{d\psi_i}{dt} &= 2a_ix_i - \sum_{j=n}^{n} a_{ji}\psi_j \quad (i=1, 2, \ldots, n). \end{aligned}\right\} \tag{3.314}$$

Ce système d'équations doit être résolu avec les conditions initiales $x_1(0), x_2(0), \ldots, x_n(0)$, ainsi que avec les conditions aux limites (finales)

$$\lim_{t\to\infty} x_1(t) = \lim_{t\to\infty} x_2(t) = \ldots = \lim_{t\to\infty} x_n(t) = 0. \tag{3.315}$$

Il faut sélectionner les $\psi_i(0)$ $(i = 1, 2, \ldots, n)$ de la sorte que soient vérifiées les conditions initiales de même que les conditions finales pour $x_i(t)$.

Il est montré dans [3.39] que les valeurs recherchées de $\psi_i(0)$ sont associées aux $x_j(0)$ par des relations linéaires

$$\psi_i(0) = \sum_{j=1}^{n} \gamma_{ij}x_j(0) \quad (i=1, 2, \ldots, n), \tag{3.316}$$

où γ_{ij} sont des constantes. Tout instant pouvant être pris pour initial, en général, dans ce problème pour n'importe quel instant

$$u^* = \frac{1}{2c} \sum_{i=1}^{n} k_i x_i, \tag{3.317}$$

où

$$k_i = \sum_{j=1}^{n} b_j \gamma_{ji}. \tag{3.318}$$

Par analogie avec (3.313) on supposait [3.28, 3.39] que dans le cas général la loi de la commande optimale s'écrivait

$$u = \begin{cases} \frac{1}{2c} \sum_{i=1}^{n} k_i x_i, & \text{si } \left| \frac{1}{2c} \sum_{i=1}^{n} k_i x_i \right| \leqslant 1. \\ 1, & \text{si } \frac{1}{2c} \sum_{i=1}^{n} k_i x_i \geqslant 1, \\ -1, & \text{si } \frac{1}{2c} \sum_{i=1}^{n} k_i x_i \leqslant -1. \end{cases} \tag{3.319}$$

Mais cette hypothèse s'est trouvée fausse [3.48]. La loi de commande (3.319) n'est strictement optimale que dans les cas particuliers, où le point représentatif, après avoir rejoint le domaine délimité par les hyperplans

$$\frac{1}{2c} \sum_{i=1}^{n} k_i x_i = +1, \quad \frac{1}{2c} \sum_{i=1}^{n} k_i x_i = -1, \tag{3.320}$$

n'en sort plus pendant son mouvement ultérieur.

Dans [3.40] et [3.41] est décrite la théorie et sont donnés les résultats de l'étude d'une commande asservie à deux paramètres de commande. Le système est voisin d'un système optimal à temps minimal. Le problème a été résolu par l'application du principe du maximum. Il semble que ces ouvrages sont les premiers à appliquer le principe du maximum aux problèmes techniques. L'objet de commande est un moteur à courant continu à excitation indépendante. Le moteur reçoit deux commandes. La tension U_1 en est une, elle est amenée à l'entrée de l'amplificateur rotatif (AR) qui alimente le circuit de l'induit du moteur. L'autre commande est la tension U_2 qui alimente la bobine d'excitation du moteur. La conduite est assurée simultanément par les deux commandes U_1 et U_2.

L'étage de puissance du système asservi est schématisé par la figure 3.26. I_a est ici l'intensité de la bobine d'excitation de l'AR

à constante de temps T_1. L'intensité de la bobine d'excitation du moteur présentant la constante de temps T_2 se note I_e. La constante de temps de la bobine transversale de l'AR est désignée par T_3, alors que la force électromotrice de l'AR est désignée par E_g. Soit R_i la résistance totale des circuits des induits de l'AR et du moteur, Ω la vitesse de l'arbre, $E_m = c\Omega I_e$ la force contre-électromotrice du moteur, I_i l'intensité de l'induit, $M_m = k_3 I_i I_e$ le couple de rotation, J le moment d'inertie réduit à l'arbre du moteur, X la position angulaire de l'arbre de sortie. Nous négligeons l'induction dans les circuits des induits, la réaction de l'induit et le couple résistant sur l'arbre du moteur. Les équations du schéma de la figure 3.26 peuvent alors s'écrire

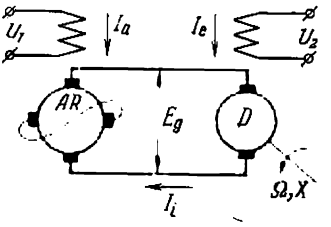


Fig. 3.26.

$$\left.\begin{aligned} &\frac{dX}{dt} = \zeta\Omega, && J\frac{d\Omega}{dt} = k_3 I_i I_e, \\ &T_1\frac{dI_a}{dt} + I_a = k_1 U_1, && T_2\frac{dI_e}{dt} + I_e = k_2 U_2, \\ &T_3\frac{dE_g}{dt} + E_g = k_4 I_a, && E_g - c\Omega I_e = I_e R_i. \end{aligned}\right\} \tag{3.321}$$

ζ, k_1, k_2, k_3, k_4 sont ici les gains constants des circuits. Si $T_1 \ll T_3$, après les transformations et la réduction au temps relatif, le système des équations de l'objet de la figure 3.26 se met sous la forme

$$\left.\begin{aligned} &\frac{dx}{dt} = \zeta_0\omega, && \frac{d\omega}{dt} = \frac{1}{1-\mu} i_e e_g - \omega i_e^2, \\ &\frac{de_g}{dt} = -\beta_1 e_g + \beta_1 u_1, && \frac{di_e}{dt} = -\beta_2 i_e + \beta_2 u_2. \end{aligned}\right\} \tag{3.322}$$

Dans ces équations nous avons noté

$$\left.\begin{aligned} &x = \frac{X}{X_{\max}}, \quad e_g = \frac{E_g}{E_{g\,\max}}, \quad i_e = \frac{I_e}{I_{e\,\max}}, \\ &\omega = \frac{\Omega}{\Omega_{\max}}, \quad u_1 = \frac{U_1}{U_{1\,\max}}, \quad u_2 = \frac{U_2}{U_{2\,\max}}, \\ &\beta_1 = \frac{T_{\text{ém}}}{T_3}, \quad \beta_2 = \frac{T_{\text{ém}}}{T_2}, \\ &T_{\text{ém}} = \frac{JR_i}{k_3 c I_{e\,\max}^2}, \quad \eta = \frac{E_{g\,\max}}{E_{g\,\text{nom}}} > 1, \\ &\qquad\qquad \mu = \frac{I_{i\,\text{nom}} R_i}{E_{g\,\text{nom}}} < 1; \end{aligned}\right\} \tag{3.323}$$

Les indices « max » et « nom » sont affectés respectivement aux valeurs maximale et nominale de la grandeur intéressée.

Dans les équations (3.222) de même que dans [3.40], nous avons gardé la notation t pour le temps sans dimension ; nous avons pris pour unité de base la constante de temps électromécanique $T_{\text{ém}}$.

La tension U_1 ne peut pas dépasser en valeur absolue la valeur maximale admissible. D'autre part, on ne peut pas diminuer excessivement la tension U_2 au risque de provoquer l'emballement du moteur pour un faible couple de charge. C'est pourquoi u_1 et u_2 sont soumises aux contraintes

$$|u_1| \leqslant 1, \ \lambda \leqslant u_2 \leqslant 1, \tag{3.324}$$

où $\lambda = \text{const} > 0$.

Considérons la réponse transitoire à l'écart initial, équivalent au saut A_0 de la consigne, où A_0 est une constante arbitraire. Introduisons de nouvelles coordonnées associées aux anciennes par les égalités

$$x_1 = A_0 - x, \quad x_2 = \frac{dx_1}{dt} = -\zeta_0\omega, \quad x_3 = -e_g, \quad x_4 = i_e. \tag{3.325}$$

Portons ces expressions dans (3.322), il vient

$$\left.\begin{aligned} \frac{dx_1}{dt} &= x_2, & \frac{dx_2}{dt} &= \alpha x_3 x_4 - x_2 x_4^2, \\ \frac{dx_3}{dt} &= -\beta_1 x_3 - \beta_1 u_1, & \frac{dx_4}{dt} &= -\beta_2 x_4 + \beta_2 u_2, \end{aligned}\right\} \tag{3.326}$$

où

$$\alpha = \frac{\eta}{1-\mu}. \tag{3.327}$$

Soient les valeurs initiales $x_i(0)$ $(i = 1, 2, 3, 4)$. Recherchons les commandes optimales $u_1(t)$ et $u_2(t)$ recalant le point représentatif à zéro en temps minimal. Le point zéro est défini par les coordonnées $(0, 0, 0, \gamma)$ où $\lambda < \gamma < 1$ est un certain nombre fixé.

Les équations (3.326) peuvent être mises sous la forme vectorielle

$$\frac{d\bar{x}}{dt} = \bar{f}(\bar{x}) + B\bar{u}, \tag{3.328}$$

où $\bar{x}$ et $\bar{f}$ sont les vecteurs dont les composantes sont données par (3.326) ; les composantes du vecteur $\bar{u}$ sont u_1 et u_2, alors que B est une matrice rectangulaire

$$B = \begin{vmatrix} 0 & 0 \\ 0 & 0 \\ -\beta_1 & 0 \\ 0 & \beta_2 \end{vmatrix}. \tag{3.329}$$

Composons le système des équations adjointes

$$\left.\begin{aligned} &\frac{d\psi_1}{dt}=0, \quad \frac{d\psi_2}{dt}=-\psi_1+x_4^2\psi_2,\\ &\frac{d\psi_3}{dt}=-\alpha x_4\psi_2+\beta_1\psi_3,\\ &\frac{d\psi_4}{dt}=(2x_2x_4-\alpha x_3)\,\psi_2+\beta_2\psi_4. \end{aligned}\right\} \tag{3.330}$$

Formons l'hamiltonien

$$H=\left\langle\bar{\psi},\ \frac{d\bar{x}}{dt}\right\rangle=\langle\bar{\psi},\ \bar{f}(\bar{x})+B\bar{u}\rangle=\langle\bar{\psi},\ \bar{f}(\bar{x})\rangle+\langle\bar{\psi},\ B\bar{u}\rangle. \tag{3.331}$$

Dans cette expression seul le second terme dépend de $\bar{u}$. Ecrivons-le sous une forme développée

$$\langle\bar{\psi},\ B\bar{u}\rangle=-\psi_3(t)\,\beta_1u_1+\psi_4(t)\beta_2u_2. \tag{3.332}$$

Il en résulte, compte tenu de (3.324), que le maximum de H a lieu sous les conditions suivantes:

$$u_1=-\operatorname{sign}\psi_3(t) \tag{3.333}$$

et

$$u_2=\begin{cases}1, & \text{si } \psi_4(t)>0,\\ \lambda, & \text{si } \psi_4(t)<0.\end{cases} \tag{3.334}$$

Ainsi, si la commande optimale existe, elle est composée de plusieurs intervalles dans chacun desquels les actions u_1 et u_2 sont maintenues à l'une de leurs valeurs limites.

Pour expliciter la loi de la commande optimale, il faut connaître le vecteur $\bar{\psi}$ en chaque point $\bar{x}$ de la trajectoire optimale. Comme nous l'avons constaté au chapitre II, dans le cas général, la résolution du problème aux limites correspondant n'est possible qu'à l'aide des calculateurs. Mais dans certains cas la synthèse d'un système optimal peut être réalisée par la méthode analytique. Considérons le cas le plus simple, quand les constantes de temps T_2 de la bobine transversale du générateur et T_3 de la bobine d'excitation du moteur sont négligeables par rapport à $T_{\text{ém}}$, constante de temps électromécanique (dans [3.40 et 3.41] sont traités des cas plus complexes). On a alors $\beta_1=\infty$ et $\beta_2=\infty$ et les équations de l'objet deviennent

$$\frac{dx_1}{dt}=x_2, \quad \frac{dx_2}{dt}=-u_2^2x_2+\alpha u_1u_2. \tag{3.335}$$

Les contraintes imposées à u_1 et u_2 s'expriment toujours par les conditions (3.324). Le système adjoint s'écrit maintenant

$$\frac{d\psi_1}{dt}=0 \quad \frac{d\psi_2}{dt}=-\psi_1+u_2^2\psi_2. \tag{3.336}$$

Formons l'hamiltonien

$$H = \langle \overline{\psi}, \overline{f}(x) \rangle = \psi_1 x_2 + \psi_2 (-u_2^2 x_2 - \alpha u_1 u_2). \tag{3.337}$$

Les solutions des équations (3.336) sont

$$\psi_1(t) = \psi_{10} = \text{const},$$

$$\psi_2(t) = \exp\left\{\int_0^t u_2^2(\tau)\,d\tau\right\}\left[\psi_{20} - \int_0^t \psi_{10} \exp\left\{-\int_0^t u_2^2(s)\,ds\right\} d\tau\right]. \tag{3.338}$$

Cette dernière expression montre que $\psi_2(t)$ ne peut changer de signe qu'une seule fois.

Les grandeurs u_2 et α de la formule (3.337) étant toujours positives, le choix de u_1 maximisant H entraîne l'expression

$$u_1 = -\operatorname{sign} \psi_2(t). \tag{3.339}$$

Pour élucider la loi de commande pour $u_2(t)$ mettons la formule (3.337) sous la forme suivante:

$$H = \psi_1 x_2 + \psi_2 \left[-x_2\left(u_2 + \frac{\alpha u_1}{2x_2}\right)^2 + \frac{\alpha^2 u_1^2}{4x_2}\right]. \tag{3.340}$$

Deux cas sont possibles.

a) $\psi_2(t) < 0$, $u_1 = +1$. Avec $x_2 > 0$ la grandeur u_2 doit prendre également sa valeur maximale, égale à l'unité. Mais si $x_2 < 0$, on voit alors de (3.340) que pour maximiser H il faut prendre

$$u_2 = \begin{cases} 1, & \text{si } \dfrac{\alpha}{2|x_2|} \geqslant 1, \\ \dfrac{\alpha}{2|x_2|}, & \text{si } \lambda \leqslant \dfrac{\alpha}{2|x_2|} \leqslant 1, \\ \lambda, & \text{si } \dfrac{\alpha}{2|x_2|} \leqslant \lambda. \end{cases} \tag{3.341}$$

b) $\psi_2(t) > 0$, $u_1 = -1$. La maximisation de H avec $x_2 > 0$ nécessite alors

$$u_2 = \begin{cases} 1, & \text{si } \dfrac{\alpha}{2x_2} \geqslant 1, \\ \dfrac{\alpha}{2x_2}, & \text{si } \lambda \leqslant \dfrac{\alpha}{2x_2} \leqslant 1, \\ \lambda, & \text{si } \dfrac{\alpha}{2x_2} \leqslant \lambda. \end{cases} \tag{3.342}$$

Mais si $x_2 < 0$, la valeur de $\left(u_2 + \dfrac{2u_1}{2x_2}\right)$ doit être la plus grande possible et pour cela il faut poser $u_2 = 1$.

Comme $\psi_2(t)$ change de signe pas plus d'une fois, $u_1(t)$ en général est composé de deux intervalles au plus, dans l'un d'eux $u_1 = +1$

et dans l'autre $u_1 = -1$. La courbe $u_2(t)$ en comporte quatre. Trois d'entre eux correspondent au premier intervalle de u_1 et le quatrième au deuxième intervalle de u_1. D'après la première ligne de (3.341),

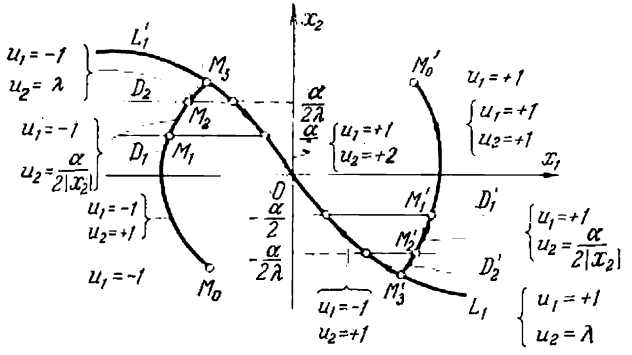


Fig. 3.27.

u_2 est maximal dans le premier intervalle. Dans le deuxième intervalle u_2 varie en permanence et parcourt le segment $[\lambda, 1]$ [cf. les deuxièmes lignes des formules (3.341) et (3.342)]. Dans le troisième intervalle $u_2 = \lambda$. Enfin, dans le dernier intervalle l'action u_2 doit prendre de nouveau sa valeur maximale, pour qu'à la fin de cet intervalle, quand $x_1 = x_2 = 0$ l'hamiltonien H devient maximal.

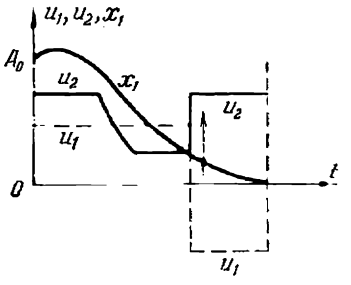


Fig. 3.28.

Les trajectoires optimales dans le plan de phase sont représentées sur la figure 3.27 et les courbes approchées $u_1(t)$, $u_2(t)$, $x(t)$ sur la figure 3.28. Seulement deux trajectoires optimales aboutissent à l'origine des coordonnées. L'une d'elles, désignée par L_1, correspond aux valeurs $u_1 = -1$, $u_2 = +1$, l'autre, notée L_1', correspond à $u_1 = +1$, $u_2 = +1$. A ces trajectoires aboutissent les trajectoires des familles L_2 et L_2'. Considérons, par exemple, la trajectoire typique $M_0M_1M_2M_3$ qui mène à L_1'. Tout le domaine situé à gauche de la courbe de commutation L_1, L_1' est relatif à $u_1 = -1$, alors que tout le domaine à droite de cette courbe est caractérisé par $u_1 = +1$. Ainsi, l'action de commande u_1 subit une seule commutation au cours du processus de régulation, ce qu'on a représenté sur la figure 3.28. Quant à u_2, la loi de variation de cette action est plus complexe. Traçons dans le domaine $u_1 = -1$ deux droites horizontales D_1

$\left(\text{droite } x_2 = \frac{\alpha}{2}\right)$ et D_2 $\left(\text{droite } x_2 = \frac{\alpha}{2\lambda}\right)$. Les droites analogues dans le domaine $u_1 = +1$ sont désignées D_1' $\left(\text{droite } x_2 = -\frac{\alpha}{2}\right)$ et D_2' $\left(\text{droite } x_2 = -\frac{\alpha}{2\lambda}\right)$.

Si le point initial M_0 se situe dans le troisième quadrant, il se déplace suivant la trajectoire optimale $M_0M_1M_2M_3O$. Sur la portion $M_0M_1M_2M_3$ de cette trajectoire, $u_1 = -1$ et sur M_3O, $u_1 = +1$. Sur la portion M_0M_1 la valeur de $u_2 = 1$, sur la portion M_1M_2, entre les droites D_1 et D_2, u_2 varie selon la loi $u_2 = \frac{\alpha}{2x_2}$. Au point M_2 la valeur de u_2 devient égale à λ et se maintient à ce niveau jusqu'au point M_3, où les deux grandeurs, u_1 et u_2 changent brusquement de valeur, u_2 devient égale à $+1$ et se maintient à ce niveau sur la portion M_3O.

Dans le cas de la trajectoire optimale $M_0'M_1'M_2'M_3'O$ issue du point M_0' situé dans le premier quadrant, la variation de u_1 et u_2 est analogue, à la différence que sur la portion $M_0'M_1'M_2'M_3'$ on a $u_1 = +1$ et entre M_3' et O $u_1 = -1$. La valeur de u_2 varie de la même façon que pour la trajectoire $M_0M_1M_2M_3O$. Les courbes approchées de variation de $x_1(t)$, $u_2(t)$ et $u_1(t)$, ce dernier étant représenté en pointillé à une échelle différente de u_2, sont montrées sur la figure 3.28. Ils correspondent à la trajectoire $M_0'M_1'M_2'M_3'O$. On distingue sur la figure 3.28 deux intervalles de la courbe $u_1(t)$ et quatre de la courbe $u_2(t)$.

CHAPITRE IV

SYSTÈMES OPTIMAUX À INFORMATION INCOMPLÈTE MAXIMALE SUR L'OBJET COMMANDÉ

§ 1. Systèmes continus à information maximale sur l'objet

Dans le chapitre I nous avons donné la définition des systèmes à information complète et incomplète sur l'objet commandé. Nous considérons que l'information sur l'opérateur F de l'objet et le critère final Q de la commande est toujours complète. Dans ce chapitre nous adoptons également que l'information sur l'état $\bar{x}$ de l'objet est aussi complète. Autrement dit, les renseignements sur l'état de l'objet B sont canalisés vers l'organe de commande A par le circuit de réaction sans perturbations. Supposons, de plus, que la consigne $\bar{x}^*$ et la perturbation $\bar{z}$ sont mesurées également sans erreur et appliquées à l'organe de commande A. L'organe A peut mesurer sans erreur et mémoriser ses valeurs de sortie $\bar{u}$.

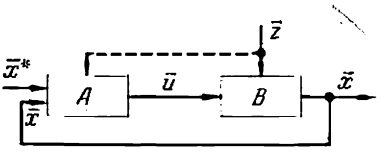


Fig. 4.1.

Considérons le système schématisé par la figure 4.1. L'organe de commande reçoit la consigne $\bar{x}^*$ et la perturbation $\bar{z}$ préalablement mesurée (flèche en pointillé). Supposons maintenant que l'une des actions $\bar{x}^*$ et $\bar{z}$ ou les deux ensemble sont des processus aléatoires. Même la connaissance parfaite de la « préhistoire » de ces processus pendant le temps τ, $-\infty < \tau \leqslant t$, où t est l'instant courant, ne rend pas possible la prévision de leurs valeurs futures. Il en résulte que l'information sur l'objet gouverné conservée dans l'organe de commande A est incomplète.

Il convient de noter que la mesure de la perturbation $\bar{z}$ peut être réalisée de deux façons.

a) *Mesure directe.* Dans ce cas le trait interrompu de la figure 4.1 doit être remplacé par un trait plein. La perturbation $\bar{z}$ est mesurée par un appareil de mesure quelconque et le résultat est transmis à l'organe de commande A. Nous considérons l'erreur de mesure négligeable.

b) *Mesure indirecte.* Dans plusieurs cas $\bar{z}$ peut s'obtenir par une voie indirecte, en mesurant les valeurs de $\bar{u}$ et de $\bar{x}$ fournies à l'organe A [4.1]. En effet, soient les équations de l'objet B de la forme

$$\frac{dx_i}{dt} = f_1(\bar{x}, \bar{u}, z) \quad (i = 1, \ldots, n), \tag{4.1}$$

où z est un scalaire. Examinons l'une des équations (4.1), par exemple la première :

$$\frac{dx_1}{dt} = f_1(\bar{x}, \bar{u}, z). \tag{4.2}$$

Soit f_1 la fonction univoque et monotone du scalaire z, quelles que soient les valeurs de $\bar{x}$ et de $\bar{u}$. En mesurant les x_i dans un intervalle infiniment petit ou en dérivant x_i dans un dérivateur parfait, on peut trouver en principe avec une erreur aussi petite que l'on veut la valeur de la dérivée $\frac{dx_i}{dt}$ et, par conséquent, la valeur de $f_1(\bar{x}, \bar{u}, z)$. Connaissant $\bar{x}$ et $\bar{u}$, on peut définir z.

Si un système sans mémoire (sans inertie) est décrit par l'équation

$$x = F(u, z), \tag{4.3}$$

où F est une fonction monotone de z, les valeurs courantes de x et de $\bar{u}$ permettent de définir également la valeur de z. Par exemple, dans le cas le plus simple

$$x = \varphi(u) + z, \tag{4.4}$$

la grandeur z est égale à $x - \varphi(u)$. Il est évident que dans de nombreux cas il est possible de mesurer plusieurs composantes de la perturbation $\bar{z}$. Ainsi nous considérons que $\bar{x}$ et $\bar{u}$ sont mesurables sans erreur et que l'opérateur F de l'objet est également bien connu et permet de définir $\bar{z}$ à partir de $\bar{x}$ et $\bar{u}$. Nous sommes alors dans le cas où le résultat d'une mesure indirecte ne se distingue en rien de celui d'une mesure directe.

Dans ce qui suit nous poserons qu'à un organe optimal s'applique le *principe d'utilisation complète de l'information* qui peut être énoncé sous la forme suivante :

Un organe de commande optimal utilise pour la commande toute l'information qui lui est fournie. Bien entendu il ne s'agit là que de l'information utile.

La formulation des problèmes de commande optimale doit être vérifiée à la lumière du principe d'utilisation complète de l'information disponible. Il se peut qu'un énoncé à première vue correct donne lieu à des inexactitudes. Ces inexactitudes peuvent être de deux types.

a) *Redondance d'information ou excès de structure.* Considérons le schéma de la figure 4.2,*a*. Un appareil spécial mesure la perturbation $\bar{z}$ et envoie le résultat à l'organe de commande A. Mais si dans le système une mesure indirecte de $\bar{z}$ est possible, la chaîne de mesure entre $\bar{z}$ et A est alors inutile. L'information qu'elle canalise est excédentaire et l'organe de mesure lui-même est superflu dans la structure générale.

b) *Utilisation incomplète de l'information.* Supposons que la perturbation $\bar{z}$ se mesure avec une erreur ou qu'un bruit aléatoire $\bar{e}$ vient entâcher le résultat de mesure lors du passage du signal par le circuit E (fig. 4.2,*b*). Supposons données les caractéristiques probabilistes de $\bar{e}$ et de $\bar{z}$. Supposons également qu'on peut mesurer indirectement et avec une précision suffisante la perturbation $\bar{z}$. Dans ce cas la chaîne de mesure et le circuit E sont superflus de même que l'information sur les caractéristiques probabilistes du bruit $\bar{e}$. Cependant ce serait une erreur de se borner aux informations sur $\bar{z}$ fournies par le circuit E et de considérer, de plus, que les renseignements dont dispose l'organe de commande A sur la perturbation $\bar{z}$ sont inexacts. Une telle formulation du problème est en contradiction avec le principe d'utilisation complète de l'information. La même contradiction survient si la chaîne de mesure directe de $\bar{z}$ est absente et si l'on admet que l'organe de commande ne dispose que d'une information a priori sur la perturbation $\bar{z}$ (par exemple, sur la densité de probabilité a priori de la variable aléatoire $\bar{z}$), alors qu'il est possible d'en réaliser une mesure indirecte précise. Si $\bar{z}$ est une variable aléatoire, sa valeur dans les systèmes de ce type devient connue avec précision peu après le démarrage du système. En remplaçant la valeur précise par une répartition probabiliste a priori on risque d'altérer sensiblement les performances du système qui se trouve loin d'être optimal.

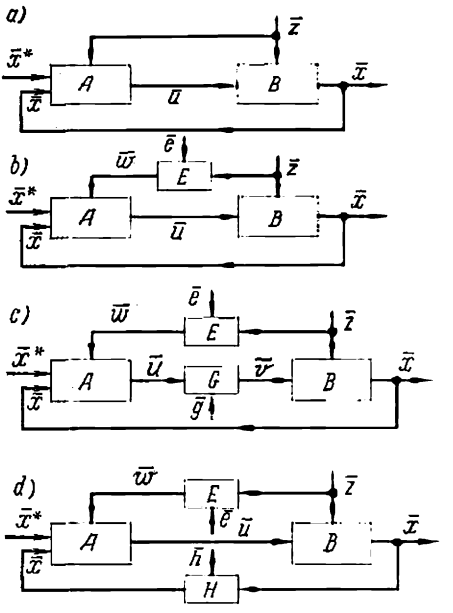


Fig. 4.2.

Dans des systèmes réels aucune grandeur ne peut être mesurée avec une précision absolue. Tous les systèmes de transmission ou de traitement de l'information donnent lieu à des bruits. C'est, par exemple, le bruit $\bar{g}$ dans le circuit G réunissant A à B (fig. 4.2,*c*) ou le bruit $\bar{h}$ dans le circuit H du circuit de réaction (fig. 4.2,*d*). Dans ce cas la mesure indirecte précise de la perturbation $\bar{z}$ est impossible et le circuit de mesure directe E peut fournir une information supplémentaire sur la perturbation. Dans chaque cas concret on évalue la valeur de cette information supplémentaire et on décide s'il est nécessaire d'introduire le circuit de mesure directe de $\bar{z}$ dans le système.

Une question se pose : est-il nécessaire pour la commande optimale de disposer dans tous les cas d'une information complète sur l'objet ? Et notamment, si la commande optimale devient beaucoup moins efficace à cause de l'absence de l'information sur la perturbation z ? Du point de vue pratique, cette question est très importante car le plus difficile est souvent d'obtenir une information sur la perturbation z ou, ce qui revient au même, sur les caractéristiques aléatoires du système gouverné. Il s'avère que dans certains cas l'information sur la perturbation z est inutile pour définir la commande optimale. Soit, par exemple, le système décrit par les équations

$$\left.\begin{aligned} \frac{dv}{dt} &= u, \\ x &= F(v, z), \end{aligned}\right\} \tag{4.5}$$

où la perturbation z est constante au cours du processus transitoire, alors que F est une fonction monotone de v ; de plus, $\frac{\partial F}{\partial v} > 0$. Adoptons une contrainte sous la forme $|u| \leqslant U$. Supposons qu'il faut réaliser un système optimal à temps minimal, où est minimisé le temps nécessaire pour éliminer l'écart $(x^* - x)$, où $x^* = \text{const}$. Dans le système décrit par les équations (4.5), la commande optimale s'écrit

$$u^* = U \operatorname{sign}(x^* - x) \tag{4.6}$$

et ne dépend pas de la valeur de z. C'est pourquoi l'information sur la perturbation z est ici excédentaire.

Pourtant, les exemples de ce genre sont très rares et dans un certain sens dégénérés. Dans la grande majorité des cas, la mise à profit de l'information sur la perturbation améliore le résultat de la commande.

Si la consigne $\bar{x}^*$ et la perturbation $\bar{z}$ se mesurent avec précision et sont des fonctions régulières du temps, un intervalle infiniment petit de temps suffit pour établir avec précision les valeurs futures

de ces fonctions. Dans ce cas nous sommes ramenés aux systèmes à information complète sur l'objet gouverné. Pourtant si $\bar{x}^*$ et $\bar{z}$ sont des processus aléatoires, l'organe de commande, comme nous l'avons déjà dit, ne dispose plus d'une information complète sur l'objet. Deux cas peuvent alors se présenter.

a) Dans le cas général on peut, en observant la « préhistoire » du processus aléatoire, accumuler une information supplémentaire qui permettra d'estimer avec assez de précision ses valeurs futures et de remplacer les caractéristiques probabilistes a priori par des caractéristiques a posteriori. Les systèmes de ce type sont rangés dans la classe des systèmes à accumulation de l'information sur l'objet. Ils sont étudiés dans les chapitres qui suivent.

b) Dans le cas particulier, lorsque $\bar{x}^*$ et $\bar{z}$ sont des processus aléatoires ou markoviens et lorsqu'on mesure avec précision leurs valeurs courantes, l'accumulation de l'information n'a pas lieu. En effet, un processus strictement aléatoire se distingue par ce que ses valeurs futures ne dépendent pas du passé. C'est pourquoi, si l'on connaît la caractéristique probabiliste a priori du processus, la connaissance de sa « préhistoire » ne fournit aucun renseignement supplémentaire sur ses valeurs futures possibles. Nous n'examinons pas le cas où cette caractéristique est inconnue et la connaissance de la « préhistoire » permet de l'évaluer et ceci avec une précision d'autant plus grande que le temps de l'observation est plus grand. Si $\bar{x}^*$ et $\bar{z}$ sont des processus markoviens, leurs valeurs futures ne sont associées qu'à leurs valeurs courantes et la connaissance de la « préhistoire » de ces processus ne donne aucune information supplémentaire. Alors, dans ce cas aussi l'accumulation de l'information n'est pas nécessaire non plus. Dans les cas qui viennent d'être traités, l'information maximale possible sur l'objet est fournie à l'organe de commande A si les mesures précises des valeurs courantes de $\bar{x}^*$ et de $\bar{z}$ sont transmises à son entrée. Il est naturel d'appeler les systèmes de ce type systèmes à information *maximale* (mais incomplète) sur l'objet.

Dans [4.2] sont décrits certains types de processus aléatoires appliqués à l'entrée des systèmes de commande automatique. Nous ne traitons dans ce qui suit que des processus strictement aléatoires et markoviens aléatoires. Plus généralement, un processus markovien peut être considéré également comme le résultat du passage d'un processus strictement aléatoire par un certain système dynamique. Soit l'équation d'un tel système à sortie $\bar{z}$ et à entrée $\bar{\xi}$

$$\frac{d\bar{z}}{dt} = \bar{\varphi}(\bar{z}, \bar{\xi}, t). \tag{4.7}$$

Les coordonnées φ_i du vecteur $\bar{\varphi}$ $(i = 1, \ldots, m)$ sont ici certaines fonctions, en général, non linéaires, de leurs arguments (qui peuvent être considérées comme dérivables).

Posons que le vecteur $\bar{\xi}$ est le vecteur d'un processus strictement aléatoire amené à l'entrée du système décrit par l'équation (4.7). A l'instant donné on ne peut conclure sur les valeurs futures de $\bar{\xi}$ que d'après la densité de probabilité a priori connue $P(\bar{\xi})$. L'observation d'un processus strictement aléatoire ne donne aucune information nouvelle. Pourtant, à un certain instant futur la valeur de $\bar{z}$ dépend non seulement des valeurs futures de $\bar{\xi}$, mais également de la position du point représentatif $\bar{z}$ dans l'espace de phase de dimension m du système (4.7) à l'instant courant. Par conséquent, la densité de probabilité de la grandeur $\bar{z}$ à l'instant futur $(t+\tau)$ dépend de la valeur de $\bar{z}$ à l'instant courant t. De plus, la connaissance de la « préhistoire », c'est-à-dire des valeurs de $\bar{z}$ aux instants passés, n'ajoute rien à l'information sur la marche future du processus $\bar{z}$. Il est donc évident que $\bar{z}$ est un processus markovien.

Si on ajoute les équations (4.7) aux équations de l'objet, on obtient un système « équivalent » qui subit l'action d'un processus strictement aléatoire $\bar{\xi}(t)$.

La différence importante entre les problèmes examinés dans ce chapitre et ceux du chapitre précédent consiste également dans la nature du critère d'optimalité. Les grandeurs d'entrée $\bar{x}^*$ et $\bar{z}$ introduites dans le système étant des processus aléatoires, la grandeur de sortie $\bar{x}$ de l'objet B et la commande $\bar{u}$ le sont aussi. Par conséquent, si l'on choisit un critère d'optimalité primaire quelconque $Q_1 = Q_1(\bar{x}^*, \bar{x}, \bar{u}, \bar{z})$, Q_1 est également une grandeur aléatoire qui change d'une expérience à l'autre d'une façon imprévisible. Pourtant, le critère d'optimalité définitif Q ne doit pas être une mesure aléatoire de la qualité du fonctionnement du système. C'est pourquoi on adopte généralement comme mesure Q soit l'espérance mathématique du critère primaire Q_1, donc $Q = M\{Q_1(\bar{x}^*, \bar{x}, \bar{u}, \bar{z})\}$, soit la probabilité du fait que le critère Q_1 sera suffisamment petit (inférieur au seuil imposé), soit encore le seuil Q tel que l'événement $Q_1 < Q$ ait la probabilité imposée.

Dans le cas général, Q_1 est une fonctionnelle des processus $\bar{x}^*$, $\bar{x}$, $\bar{u}$, $\bar{z}$ dans un intervalle de temps fini ou infini. Dans ce qui suit nous examinons un problème plus restreint, où l'intervalle de temps est fini. Soit $t_0 \leqslant t \leqslant T$, de plus, les valeurs de t_0 et T sont fixées.

Soit l'équation simple de premier ordre

$$\frac{dx}{dt} = f(x, u, t) + \xi, \tag{4.8}$$

où ξ est un processus strictement aléatoire. Figurons d'abord le processus $\xi(t)$ sous la forme d'une suite discrète des grandeurs gaussien-

nes aléatoires indépendantes de moyenne $m(x, t)$ et de variance $\sigma^2(x, t)/\Delta t$. Soit Δt l'intervalle entre les échantillons.

Pour $\Delta t \to 0$ on obtient à la limite un processus strictement aléatoire à variance infinie. Il convient de noter qu'un bruit blanc, processus aléatoire stationnaire de densité spectrale constante $S(\omega) = S_0$ (cf. chapitre II), a également une variance infiniment grande. En effet, en substituant S_0 à $S_x(\omega)$ dans la formule (2.94) on obtient la variance du bruit blanc sous la forme

$$K_x(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_0 \, d\omega \to \infty. \tag{4.9}$$

Le processus $\xi(t)$ de moyenne $m(x, t)$ et de variance $\sigma^2(x, t)/\Delta t$, décrit plus haut, est non stationnaire puisque ses caractéristiques dépendent du temps t. Par conséquent, il peut être considéré comme une généralisation du bruit blanc normal.

Pour de petits Δt, (4.8) entraîne

$$\Delta x = f\Delta t + \xi\Delta t. \tag{4.10}$$

La répartition de la grandeur $\xi\Delta t$ est évidemment aussi normale, sa moyenne est $m\Delta t$ et sa variance $\frac{\sigma^2}{\Delta t}(\Delta t)^2 = \sigma^2\Delta t$.

Ainsi

$$P(\xi\Delta t) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \exp\left\{-\frac{(\xi\Delta t - m\Delta t)^2}{2\sigma^2\Delta t}\right\}. \tag{4.11}$$

La densité de probabilité conditionnelle de l'accroissement Δx, x et t étant fixés, s'obtient en remplaçant dans (4.11) $\xi\Delta t$ par son expression tirée de (4.10):

$$P(\Delta x \mid x, t) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \exp\left\{-\frac{(\Delta x - f\Delta t - m\Delta t)^2}{2\sigma^2\Delta t}\right\}. \tag{4.12}$$

Ainsi la densité de probabilité de l'accroissement Δx dépend de x et de t, puisque f, m et σ^2 en dépendent. Par conséquent le processus $x(t)$ est markovien.

Soit maintenant le critère primaire

$$Q_1 = \int_{t_0}^{T} G(x, u, t)\, dt. \tag{4.13}$$

Le critère statistique Q se met alors sous la forme

$$Q = M\{Q_1\} = M\left\{\int_{t_0}^{T} G(x, u, t)\, dt\right\}. \tag{4.14}$$

En effet, $x(t)$ est un processus aléatoire. Par conséquent l'intégrale Q_1 est également une grandeur aléatoire et pour l'obtenir il faut rechercher l'espérance mathématique de la grandeur Q_1.

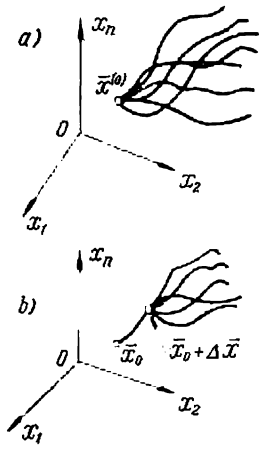


Fig. 4.3.

L'interprétation physique de la grandeur Q peut être la suivante: même si la commande choisie $u(t)$ n'est pas aléatoire, le processus $x(t)$ est quand même aléatoire. C'est pourquoi les trajectoires dans l'espace de phase, dans l'intervalle de temps $t_0 < t < T$, seront différentes pour des expériences différentes. L'« éventail » des trajectoires possibles est montré sur la figure 4.3, *a*. Si l'on mesure $\int_{t_0}^{T} G\,dt$ sur chacune de ces trajectoires possibles et si l'on prend ensuite la moyenne arithmétique, on obtient

$$Q = M\left\{\int_{t_0}^{T} G\,dt\right\}. \tag{4.15}$$

Pour définir la commande optimale $u^*(t)$ considérons l'intégrale (4.13) et découpons-la en deux intégrales. Si Δt est petit, on peut représenter cette intégrale sous la forme suivante:

$$\int_{t_0}^{T} G(x, u, t)\,dt = \int_{t_0}^{t_0+\Delta t} G(x, u, t)\,dt + \int_{t_0+\Delta t}^{T} G(x, u, t)\,dt \cong$$

$$\cong G(x_0, u_0, t_0)\,\Delta t + \int_{t_0+\Delta t}^{T} G(x, u, t)\,dt. \tag{4.16}$$

Le premier terme est transformé ici aux infiniment petits du deuxième ordre près. Supposons pour le moment que pour une certaine valeur u_0 fixée (fig. 4.3,*b*) un déplacement *donné* Δx s'effectue en un temps Δt. Considérons par la suite que lorsque $t > t_0 + \Delta t$, on a une commande optimale. L'espérance mathématique conditionnelle M' de l'intégrale (4.16) sous la condition que Δx soit donné s'écrit alors

$$M'\left\{\int_{t_0}^{T} G(x, u, t)\,dt\right\} = G(x_0, u_0, t_0)\,\Delta t +$$

$$+ \min_{\substack{u\in\Omega(u)\\ t_0+\Delta t\leq t<T}} M\left\{\int_{t_0+\Delta t}^{T} G(x, u, t)\,dt\right\}. \tag{4.17}$$

Les indices sous le signe « min » signifient que la minimisation se fait par rapport aux commandes admissibles quelconques $u(t)$ dans l'intervalle de $(t_0 + \Delta t)$ à T.

Notons Q^* la valeur minimale de Q. Elle dépend de x_0, valeur de x à l'instant $t = t_0$

$$Q^*(x_0, t_0) = \min_{\substack{u \in \Omega(u) \\ t_0 \leqslant t < T}} M\left\{\int_{t_0}^{T} G(x, u, t)\, dt\right\}. \tag{4.18}$$

Avec cette notation (4.17) s'écrit

$$M'\left\{\int_{t_0}^{T} G(x, u, t)\, dt\right\} = G(x_0, u_0, t_0)\,\Delta t + Q^*(x_0 + \Delta x, t_0 + \Delta t). \tag{4.19}$$

Mais en fait l'accroissement Δx est *aléatoire*. C'est pourquoi la valeur réelle de $Q(x_0, t_0)$ doit être l'espérance mathématique de l'expression (4.19), la moyenne $M_{\Delta x}$ étant prise sur toutes les valeurs possibles de Δx. On a

$$\begin{aligned} Q(x_0, t_0) &= M_{\Delta x}\{G(x_0, u_0, t_0)\,\Delta t + Q^*(x_0 + \Delta x, t_0 + \Delta t)\} = \\ &= G(x_0, u_0, t_0)\,\Delta t + M_{\Delta x}\{Q^*(x_0 + \Delta x, t_0 + \Delta t)\}. \end{aligned} \tag{4.20}$$

La valeur minimale Q^* de la grandeur Q s'obtient pour $u_0 = (u)_{t=t_0}$

$$Q^*(x_0, t_0) = \min_{u_0 \in \Omega(u)} \{G(x_0, u_0, t_0)\,\Delta t + M_{\Delta x}\{Q^*(x_0 + \Delta x, t_0 + \Delta t)\}\} \tag{4.21}$$

ou, sous une forme développée,

$$\begin{aligned} Q^*(x_0, t_0) = \min_{u_0 \in \Omega(u)} \Big\{ & G(x_0, u_0, t_0)\,\Delta t + \\ & + \int_{\Omega(\Delta x)} Q^*(x_0 + \Delta x, t_0 + \Delta t)\, P(\Delta x \mid x_0, t_0)\, d\Omega(\Delta x)\Big\}, \end{aligned} \tag{4.22}$$

où $P(\Delta x \mid x_0, t_0)$ est la densité de probabilité conditionnelle donnée par (4.12). Non seulement le premier, mais encore le deuxième terme dans l'accolade (4.22) dépend de u_0 du fait que $P(\Delta x \mid x_0, t_0)$ en dépend par l'intermédiaire de f [cf. (4.8) et (4.12)]. Ici $\Omega(\Delta x)$ est le domaine de toutes les valeurs possibles de Δx et $d\Omega(\Delta x)$ est son élément infiniment petit.

En général, on peut écrire dans l'équation (4.22) non pas x_0, t_0, mais x, t, car n'importe quel point courant peut être considéré comme initial. Alors, au lieu de u_0 nous figurerons dans cette équation $u = u(t)$, c'est-à-dire une valeur courante. La formule (4.20) devient

$$\begin{aligned} Q^*(x, t) = \min_{u(t) \in \Omega(u)} \Big\{ & G(x, u, t)\,\Delta t + \\ & + \int_{\Omega(\Delta t)} Q^*(x + \Delta x, t + \Delta t)\, P(\Delta x \mid x, t)\, d\Omega(\Delta x)\Big\}, \end{aligned} \tag{4.23}$$

Notons que P est une densité de probabilité et, par conséquent,

$$\int\limits_{\Omega(\Delta x)} P(\Delta x \mid x, t)\, d\Omega(\Delta x) = 1. \tag{4.24}$$

Pour passer de la relation intégrale (4.23) à l'équation différentielle aux dérivées partielles on procède de même qu'en physique, dans la théorie de la diffusion, ou en mathématiques, dans la théorie des processus markoviens [4.23, 4.24, 2.1].

Décomposons Q^* sous le signe somme en série en nous bornant aux termes du deuxième ordre

$$Q^*(x+\Delta x, t+\Delta t) = Q^*(x, t) + \frac{\partial Q^*}{\partial x}\Delta x + \frac{\partial Q^*}{\partial t}\Delta t + \frac{(\Delta x)^2}{2!}\frac{\partial^2 Q^*}{\partial x^2} + \\ + \frac{(\Delta t)^2}{2!}\frac{\partial^2 Q^*}{\partial t^2} + \Delta x \Delta t \frac{\partial^2 Q^*}{\partial x\, \partial t} + \ldots \tag{4.25}$$

Portons (4.25) dans (4.23). Compte tenu de (4.24), on a

$$Q^*(x, t) = \min_{u(t)\in\Omega(u)} \Big\{ G(x, u, t)\Delta t + Q^*(x, t) + \\ + \frac{\partial Q^*}{\partial x}\int\limits_{\Omega(\Delta x)} \Delta x P(\Delta x \mid x, t)\, d\Omega(\Delta x) + \frac{\partial Q^*}{\partial t}\Delta t + \\ + \frac{1}{2}\frac{\partial^2 Q^*}{\partial x^2}\int\limits_{\Omega(\Delta x)} (\Delta x)^2 P(\Delta x \mid x, t)\, d\Omega(\Delta x) + \frac{(\Delta t)^2}{2}\frac{\partial^2 Q^*}{\partial t^2} + \\ + \Delta t \frac{\partial^2 Q^*}{\partial x\, \partial t}\int\limits_{\Omega(\Delta x)} \Delta x P(\Delta x \mid x, t)\, d\Omega(\Delta x) + \ldots \Big\}. \tag{4.26}$$

Remarquant que les espérances mathématiques de Δx et de Δx^2 sont respectivement

$$\left.\begin{aligned} &\int\limits_{\Omega(\Delta x)} \Delta x P(\Delta x \mid x, t)\, d\Omega(\Delta x) = f\Delta t + m\Delta t, \\ &\int\limits_{\Omega(\Delta x)} (\Delta x)^2 P(\Delta x \mid x, t)\, d\Omega(\Delta x) = \sigma^2\Delta t + (m\Delta t)^2 \cong \sigma^2\Delta t, \end{aligned}\right\} \tag{4.27}$$

nous obtenons à partir de (4.27), après des simplifications que nous omettons ici, à la limite pour $\Delta t \to 0$

$$-\frac{\partial Q^*}{\partial t} = \min_{u(t)\in\Omega(u)} \Big\{ G(x, u, t) + (f+m)\frac{\partial Q^*}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 Q^*}{\partial x^2} \Big\}. \tag{4.28}$$

Les autres termes disparaissent, leur ordre étant $(\Delta t)^2$.

En recherchant la solution de l'équation (4.28) on peut trouver Q^* et en même temps la commande optimale $u^*(t)$. Cette méthode se généralise sans aucune modification (cf. [4.37]) aux systèmes d'ordre

plus élevé d'équations

$$\frac{dx_i}{dt} = f_i(x_1, \ldots, x_n) + \xi \cdot \delta_{ij} \quad (i = 1, \ldots, j, \ldots, n), \tag{4.29}$$

où j est un nombre constant de l'intervalle $1 \leqslant j \leqslant n$ et δ_{ij} est le symbole de Kronecker

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{4.30}$$

Soit ξ un processus strictement aléatoire de moyenne m et de variance $\frac{\sigma_2}{\Delta t}$. L'équation aux dérivées partielles qui se déduit de même que dans ce qui précède s'écrit dans ce cas

$$-\frac{\partial Q^*}{\partial t} = \min_{u \in \Omega(u)} \left\{ G(\bar{x}, u, t) + \sum_{i=1}^{n} (m \cdot \delta_{ij} + f_i) \frac{\partial Q^*}{\partial x_i} + \frac{\sigma^2}{2} \frac{\partial^2 Q^*}{\partial x_j^2} \right\}. \tag{4.31}$$

Aux équations (4.29) on peut ajouter également l'équation du type (4.5) qui décrit l'obtention d'un processus markovien à partir d'un processus strictement aléatoire. L'ensemble des équations (4.29) ainsi complété décrit un système « équivalent ».

Considérons l'exemple élémentaire du système décrit par l'équation

$$\frac{\partial x_1}{\partial t} = u. \tag{4.32}$$

La consigne, notée x_2, est un processus markovien obtenu en sortie d'un circuit à inertie recevant à l'entrée un processus strictement aléatoire ξ de moyenne $m = 0$

$$\frac{dx^2}{dt} = -x_2 + \xi. \tag{4.33}$$

Les équations (4.32) et (4.33) peuvent être considérées comme équations d'un certain système « équivalent » qui reçoit à l'entrée l'action ξ. Soit le critère d'optimalité

$$Q_1 = \int_{t_0}^{T} [(x_1 - x_2)^2 + u^2]\, dt. \tag{4.34}$$

L'expression (4.34) montre que dans le cas idéal il faut que $x_1 = x_2$, c'est-à-dire que la grandeur de sortie x_1 du système soit égale à la consigne x_2. Soit, ensuite, $m = 0$. L'équation (4.31) s'écrit alors

$$-\frac{\partial Q^*}{\partial t} = \min_{u} \left\{ (x_1 - x_2)^2 + u^2 + u \frac{\partial Q^*}{\partial x_1} - x_2 \frac{\partial Q^*}{\partial x_2} + \frac{\sigma^2}{2} \frac{\partial^2 Q^*}{\partial x_2^2} \right\}. \tag{4.35}$$

Si la grandeur u ne subit aucune contrainte, on minimise sans peine le second membre en le dérivant par rapport à u et en égalant à zéro

la dérivée

$$2u + \frac{\partial Q^*}{\partial x_1} = 0. \tag{4.36}$$

On en tire la valeur optimale de u^*

$$u^* = -\frac{1}{2}\frac{\partial Q^*}{\partial x_1}. \tag{4.37}$$

Si l'on porte (4.37) dans (4.35), il vient

$$-\frac{\partial Q^*}{\partial t} = (x_1 - x_2)^2 - \frac{1}{4}\left(\frac{\partial Q^*}{\partial x_1}\right)^2 - x_2\frac{\partial Q^*}{\partial x_2} + \frac{\sigma^2}{2}\frac{\partial^2 Q^*}{\partial x_2^2}. \tag{4.38}$$

Dans cette équation il est commode de passer au temps « rétrograde » τ en posant $\tau = T - t$. Alors, au lieu de (4.38), on obtient

$$\frac{\partial Q^*}{\partial \tau} = (x_1 - x_2)^2 - \frac{1}{4}\left(\frac{\partial Q^*}{\partial x_1}\right)^2 - x_2\frac{\partial Q^*}{\partial x_2} + \frac{\sigma^2}{2}\frac{\partial^2 Q^*}{\partial x_2^2}. \tag{4.39}$$

Les conditions aux limites sont naturellement les suivantes :

$$\left.\begin{aligned} &Q^*(\bar{x}, \tau = 0) = 0 \text{ pour tout } \bar{x}, \\ &Q^*(\bar{x}, \tau) \to \infty \quad \text{pour } |\bar{x}| \to \infty, \end{aligned}\right\} \tag{4.40}$$

$|\bar{x}|$ étant ici le module du vecteur $\bar{x}$ dans le plan de phase. En effet, avec $\tau = 0$, on a $t_0 = t = T$ et l'intégrale (4.34) s'annule.

La solution de l'équation (4.39) peut être recherchée, de même que dans le chapitre III, sous forme de la série

$$Q^*(\bar{x}, \tau) = k_0(\tau) + \sum k_i(\tau)\, x_i + \\ + \sum\sum k_{ij}(\tau)\, x_i x_j + \sum\sum\sum k_{ijm}(\tau)\, x_i x_j x_m. \tag{4.41}$$

Les fonctions k_0, k_i, k_{ij}, k_{ijm} ne dépendent ici que du temps. En portant (4.41) dans (4.39) et en égalant les coefficients des termes respectifs des premier et second membres de l'égalité, on peut établir que seuls $k_0(\tau)$ et $k_{ij}(\tau)$ diffèrent de zéro, on peut choisir en outre $k_{ij} = k_{ji}$. On obtient ainsi des équations différentielles ordinaires pour les fonctions $k_0(\tau)$ et $k_{ij}(\tau)$ de même que dans le chapitre III. Ces équations doivent être résolues pour les conditions initiales $k_0(0) = k_{ij}(0) = 0$ qui découlent de (4.40). Dans [4.3] sont construites les courbes des fonctions k_0, k_{11}, k_{12}, k_{22} du problème considéré ; par ailleurs, (4.41) montre que

$$Q^*(x_1, x_2, \tau) = k_0(\tau) + k_{11}(\tau)\, x_1^2 + 2k_{12}(\tau)\, x_1 x_2 + k_{22}(\tau)\, x_2^2. \tag{4.42}$$

Ces courbes sont reproduites sur la figure 4.4.

Les formules (4.37) et (4.42) donnent la loi de commande optimale

$$u^* = -k_{11}(\tau)\, x_1 - k_{12}(\tau)\, x_2. \tag{4.43}$$

Le schéma fonctionnel de l'organe de commande optimal A construit d'après l'équation (4.43) est représenté sur la figure 4.5. Le bloc D est construit d'après l'équation (4.33). Recevant à l'entrée le processus ξ strictement aléatoire, le bloc délivre à sa sortie le processus markovien x_2 qui représente la consigne pour l'organe de commande A. Bien entendu, dans un système réel le bloc D n'existe pas, il ne figure sur le dessin que pour rendre compte de la structure du processus markovien x_2. Le signe « $\int$ » dans le schéma est employé pour figurer le circuit intégrateur. Le bloc D peut être considéré comme une partie du système « équivalent » qui comporte les blocs B et D.

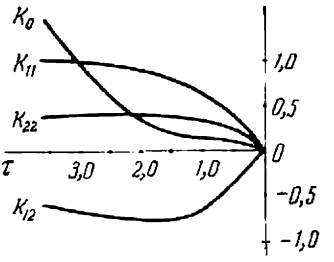


Fig. 4.4.

L'objet B est un intégrateur à grandeur de sortie x_1 [cf. (4.32)]. La commande u est élaborée dans l'organe de commande d'après l'expression (4.43) sous la forme d'une somme de signaux de sortie

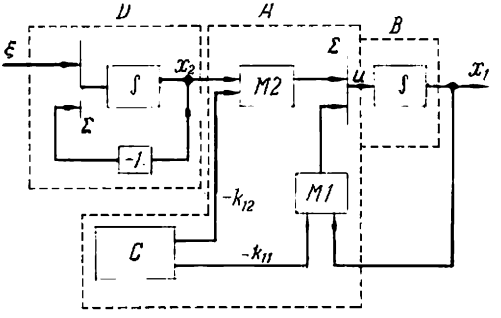


Fig. 4.5.

de deux multiplieurs $M1$ et $M2$. Les coefficients $-k_{11}(\tau)$ et $-k_{12}(\tau)$ amenés aux entrées des multiplieurs sont fournis par le calculateur C où ils sont élaborés conformément aux courbes représentées sur la figure 4.4. Il faut tenir compte du fait que $\tau = T - t$, où t est le temps réel qui varie de t_0 à T.

D'une manière analogue on peut définir la commande en temps minimal. Soit l'équation du système d'ordre n; les équations du système « équivalent » s'écrivent

$$\frac{dx_i}{dt} = f_i(\bar{x}, \bar{u}, \bar{\xi}) \quad (i = 1, \ldots, n+m), \tag{4.44}$$

où $\bar{\xi}$ est un vecteur strictement aléatoire à densité de probabilité imposée. Notons que les coordonnées x_i du système « équivalent » comptent également les coordonnées de la perturbation z_j produite par le bloc de dimension m, et que l'équation (4.44) comporte le processus strictement aléatoire ξ. Recherchons la commande optimale $\bar{u}$ qui recale en temps minimal le point représentatif $\bar{x}$ de l'espace de phase de dimension $(n + m)$ du système « équivalent », depuis la position initiale $\bar{x}^{(0)}$ à l'origine de l'espace de phase, c'est-à-dire dans le sous-domaine $x_1 = x_2 = \ldots = x_n$ de l'espace de dimension $(n + m)$. Comme $\bar{\xi}$ est un processus aléatoire, le temps T de transition est également une grandeur aléatoire et on peut prendre comme critère d'optimalité l'espérance mathématique

$$Q = M\{T\}. \tag{4.45}$$

Ensuite, au lieu du recalage du point représentatif à l'origine des coordonnées, il est opportun de considérer qu'il atteint un certain voisinage de l'origine, décrit par exemple par l'inégalité

$$\sum_{i=1}^{n} x_i^2 \leqslant \delta^2, \tag{4.46}$$

où δ est suffisamment petit.

Examinons deux positions voisines du point représentatif : $\bar{x}$ à l'instant t et $\bar{x} + \Delta\bar{x}$ à l'instant $t + \Delta t$, où Δt est petit. Soit $Q^*(\bar{x})$ la valeur minimale de l'espérance mathématique du temps de transition T si le point représentatif part de la position $\bar{x}$. De même, $Q^*(\bar{x} + \Delta\bar{x})$ est le temps de transition minimal lors du départ du point $\bar{x} + \Delta\bar{x}$, alors la somme

$$\Delta t + Q^*(\bar{x} + \Delta\bar{x}) \tag{4.47}$$

donne l'espérance mathématique du temps T, si dans une direction quelconque on fait le premier pas $\Delta\bar{x}$ *défini*, tandis que tous les autres pas sont régis par une stratégie optimale. L'existence d'une perturbation aléatoire $\bar{\xi}$ rendant en fait aléatoire la grandeur $\Delta\bar{x}$, la grandeur définie par l'expression (4.47) est également aléatoire et l'espérance mathématique $Q(\bar{x})$ de cette grandeur s'obtient en prenant la moyenne sur l'ensemble de toutes les valeurs possibles de $\Delta\bar{x}$, ce qui se note par l'opération $M_{\Delta\bar{x}}$:

$$Q(\bar{x}) = M_{\Delta\bar{x}}\{\Delta t + Q^*(\bar{x} + \Delta\bar{x})\}. \tag{4.48}$$

La commande optimale $\bar{u}^*(t)$ à l'instant t s'obtient si l'on connaît à cet instant le minimum de $Q(\bar{x})$ par rapport à la commande $\bar{u}$.

Ce minimum est égal à $Q^*(\bar{x})$. Ainsi

$$Q^*(\bar{x}) = \min_{\bar{u}\in\Omega(\bar{u})} Q(\bar{x}) = \min_{\bar{u}\in\Omega(\bar{u})} M_{\Delta\bar{x}}\{\Delta t + Q^*(\bar{x}+\Delta\bar{x})\} =$$
$$= \Delta t + \min_{\bar{u}\in\Omega(\bar{u})} M_{\Delta\bar{x}}\{Q^*(\bar{x}+\Delta\bar{x})\}. \qquad (4.49)$$

Si l'on désigne par $P(\Delta\bar{x} \mid \bar{u}, \bar{x})$ la densité de probabilité conditionnelle de $\Delta\bar{x}$, $\bar{u}$ et $\bar{x}$ étant fixés, l'égalité (4.49) se met alors sous la forme développée suivante :

$$Q^*(\bar{x}) = \Delta t + \min_{\bar{u}\in\Omega(\bar{u})} \int_{\Omega(\Delta\bar{x})} Q^*(\bar{x}+\Delta\bar{x})\, P(\Delta\bar{x} \mid \bar{u}, \bar{x})\, d\Omega(\Delta\bar{x}). \qquad (4.50)$$

De cette égalité on passe sans peine à l'équation différentielle aux dérivées partielles pour $Q^*(\bar{x})$. Développons $Q^*(\bar{x}+\Delta\bar{x})$ en série en négligeant les infiniment petits d'ordre supérieur au deuxième

$$Q^*(\bar{x}+\Delta\bar{x}) = Q^*(\bar{x}) +$$
$$+ \sum_{i=1}^{n+m} \frac{\partial Q^*}{\partial x_i}\Delta x_i + \sum_{i,j=1}^{n+m} \frac{\partial^2 Q^*}{\partial x_i \partial x_j}\Delta x_i \Delta x_j + \ldots \qquad (4.51)$$

En portant (4.51) dans (4.50) et compte tenu de ce que

$$\int_{\Omega(\Delta\bar{x})} P(\Delta\bar{x} \mid \bar{u}, \bar{x})\, d\Omega(\Delta\bar{x}) = 1, \qquad (4.52)$$

on obtient l'égalité suivante :

$$Q^*(\bar{x}) = \Delta t + \min_{\bar{u}\in\Omega(\bar{u})} \Big\{ Q^*(\bar{x}) +$$
$$+ \sum_{i=1}^{n+m} \frac{\partial Q^*}{\partial x_i} \int_{\Omega(\Delta\bar{x})} \Delta x_i P(\Delta\bar{x} \mid \bar{u}, \bar{x})\, d\Omega(\Delta\bar{x}) +$$
$$+ \sum_{i,j+1}^{n+m} \frac{\partial^2 Q^*}{\partial x_i \partial x_j} \int_{\Omega(\Delta\bar{x})} \Delta x_i \Delta x_j P(\Delta\bar{x} \mid \bar{u}, \bar{x})\, d\Omega(\Delta\bar{x}) \Big\}. \qquad (4.53)$$

Supprimons $Q^*(\bar{x})$ dans deux membres de l'égalité, divisons les deux membres par Δt et introduisons les notations

$$\left.\begin{aligned} &\lim_{\Delta t\to 0} \frac{1}{\Delta t} \int_{\Omega(\Delta\bar{x})} \Delta x_i P(\Delta\bar{x} \mid \bar{u}, \bar{x})\, d\Omega(\Delta\bar{x}) = a_i(\bar{u}, \bar{x}), \\ &\lim_{\Delta t\to 0} \frac{1}{\Delta t} \int_{\Omega(\Delta\bar{x})} \Delta x_i \Delta x_j P(\Delta\bar{x} \mid \bar{u}, \bar{x})\, d\Omega(\Delta\bar{x}) = b_{ij}(\bar{u}, \bar{x}). \end{aligned}\right\} \qquad (4.54)$$

Il est clair que a_i sont les valeurs moyennes des vitesses de variation de Δx_i, c'est-à-dire les valeurs moyennes des $\Delta x_i/\Delta t$, et que b_{ij} sont les vitesses moyennes de variation des produits des grandeurs Δx_i et Δx_j. Avec $i = j$, la grandeur b_{ij} est la vitesse moyenne de variation du carré de Δx_i, c'est-à-dire la valeur moyenne des $\frac{(\Delta x_i)^2}{\Delta t}$. Avec les notations de (4.54) l'équation (4.53) s'écrit

$$0 = 1 + \min_{\bar{u} \in \Omega(\bar{u})} \left\{ \sum_{i=1}^{n+m} \frac{\partial Q^*}{\partial x_i} a_i (\bar{u}, \bar{x}) + \sum_{i,\, j=1}^{n+m} \frac{\partial^2 Q^*}{\partial x_i \, \partial x_j} b_{ij} (\bar{u}, \bar{x}) \right\}. \quad (4.55)$$

C'est justement l'équation différentielle aux dérivées partielles recherchée dont la solution doit vérifier les conditions aux limites [cf. (4.46)]

$$Q^* (\bar{x}) = 0 \text{ avec } \sum_{i=1}^{n} x_i^2 = \delta^2 \text{ et pour tout } x_{n+j} \quad (j = 1, \ldots, m). \quad (4.56)$$

Considérons le cas particulier d'une action unique u. Les équations (4.44) du système « équivalent » sont de la forme

$$\frac{dx_i}{dt} = f_i (\bar{x}, \bar{\xi}) + \beta_i (\bar{\xi}) u \quad (i = 1, \ldots, n + m). \quad (4.57)$$

Parmi les coordonnées x_i il y a ici également les coordonnées de la perturbation, qui figure sous la forme de sortie d'un système dynamique recevant à l'entrée un processus strictement aléatoire $\bar{\xi}$. On peut alors, avec une approximation poussée aux infiniment petits de premier ordre, écrire l'accroissement Δx_i

$$\Delta x_i = f_i (\bar{x}, \bar{\xi}) \Delta t + \beta_i (\bar{\xi}) u \Delta t \quad (i = 1, \ldots, n + m). \quad (4.58)$$

On en tire la valeur moyenne de Δx_i

$$M \{\Delta x_i\} = a_i (\bar{u}, \bar{x}) \Delta t = \Delta t \, [M \{f_i (\bar{x}, \bar{\xi})\} + u M \{\beta_i (\bar{\xi})\}], \quad (4.59)$$

et la valeur moyenne du produit des accroissements

$$\begin{aligned} M \{\Delta x_i \, \Delta x_j\} &= b_{ij} (\bar{u}, \bar{x}) \Delta t = \\ &= (\Delta t)^2 M \{[f_i (\bar{x}, \bar{\xi}) + u\beta_i (\bar{\xi})] [f_j (\bar{x}, \bar{\xi}) + u\beta_j (\bar{\xi})]\} = \\ &= (\Delta t)^2 (M [f_i (\bar{x}, \bar{\xi}) f_j (\bar{x}, \bar{\xi})] + u \{M [\beta_i (\bar{\xi}) f_j (\bar{x}, \bar{\xi})] + \\ &\quad + M [\beta_j (\bar{\xi}) f_i (\bar{x}, \bar{\xi})]\} + u^2 M [\beta_i (\bar{\xi}) \beta_j (\bar{\xi})]). \quad (4.60) \end{aligned}$$

Si les fonctions de corrélation du second membre de l'égalité (4.60) sont finies, ce qui peut ne pas avoir lieu comme nous l'avons montré

plus haut, les b_{ij} sont infiniment petits et, par conséquent, disparaissent de l'équation (4.55). En se bornant à ce cas et en portant a_i tirés de (4.59) dans (4.55) on aboutit à l'équation

$$0 = 1 + \min_{u \in \Omega(u)} \sum_{i=1}^{n+m} \frac{\partial Q^*}{\partial x_i} [M\{f_i(\bar{x}, \bar{\xi})\} + uM\{\beta_i(\bar{\xi})\}]. \quad (4.61)$$

Si u est assujetti à la condition

$$|u| \leqslant 1, \quad (4.62)$$

le second membre de (4.61) est minimisé par rapport à u lorsqu

$$u = u^* = -\operatorname{sign} \sum_{i=1}^{n+m} \frac{\partial Q^*}{\partial x_i} M\{\beta_i(\bar{\xi})\}. \quad (4.63)$$

C'est la condition d'une commande optimale. On en tire que dans le type de systèmes considéré la valeur optimale u^* repose toujours sur la frontière du domaine admissible et de ce fait est égale à ± 1.

En portant (4.63) dans (4.61), il vient

$$0 = 1 + \sum_{i=1}^{n+m} \frac{\partial Q^*}{\partial x_i} M\{f_i(\bar{x}, \bar{\xi})\} - \left| \sum_{i=1}^{n+m} \frac{\partial Q^*}{\partial x_i} M\{\beta_i(\bar{\xi})\} \right|. \quad (4.64)$$

En résolvant cette équation pour les conditions aux limites (4.56) et en portant la solution obtenue dans (4.63) on peut trouver la commande optimale u^* sous une forme explicite; il se trouve de plus que c'est une fonction des valeurs courantes de x_i ($i = 1, \ldots, n + m$) dont font partie, comme nous l'avons dit dans ce qui précède, les perturbations z_j, si seulement elles constituent les sorties des circuits du système « équivalent ».

§ 2. Systèmes continus à états discrets et systèmes discrets à information maximale sur l'objet

La procédure de résolution des problèmes relatifs aux systèmes continus à état discret, c'est-à-dire aux systèmes échantillonnés dans le temps, coïncide en principe avec celle exposée au paragraphe précédent. La seule différence consiste dans ce que les équations différentielles de l'objet sont remplacées par des équations aux différences finies; de plus, après l'obtention de la relation intégrale pour Q^*, on ne passe pas à l'équation différentielle aux dérivées partielles. Illustrons cette procédure sur l'exemple d'un système

du premier ordre très simple décrit par l'équation

$$x_{k+1} = \alpha x_k + u_k + \xi_k, \tag{4.65}$$

où $\alpha = \text{const}$ et $x_k = x(k\Delta t)$, $u_k = u(k\Delta t)$, $\xi_k = \xi(k\Delta t)$. Le processus dure N cycles ($k = 0, 1, \ldots, N$), où N est imposé, alors que la durée d'un cycle est égale à Δt. Soit ξ_k le processus discret gaussien strictement aléatoire de moyenne nulle et de variance σ^2. Etant donné que (4.65) entraîne

$$\xi_k = x_{k+1} - \alpha x_k - u_k, \tag{4.66}$$

la densité de probabilité conditionnelle de x_{k+1} est définie par l'expression

$$P(x_{k+1} \mid x_k, u_k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x_{k+1} - \alpha x_k - u_k)^2}{2\sigma^2}\right\}. \tag{4.67}$$

Supposons que le critère d'optimalité s'écrit

$$Q = M\left\{\sum_{i=0}^{N} (\lambda^2 x_i^2 + u_i^2)\right\}, \tag{4.68}$$

où $\lambda^2 = \text{const}$. Il faut choisir une suite optimale $u_0, u_1, \ldots, u_N$ telle qu'elle vérifie la condition $Q = \min$. Les grandeurs u_k ne sont soumises à aucune contrainte.

L'équation intégrale pour ce problème s'obtient dans la forme de l'équation (4.22) et par des raisonnements analogues. La différence consiste dans le fait que $G\Delta t$ est remplacé par $(\lambda^2 x_k^2 + u_k^2)$ et $Q^*(x, t)$ est noté $Q^*(x_k)$, puisque dans (4.65) et (4.68) le temps t n'est pas présent explicitement. D'autre part, $Q^*(x + \Delta x, t + \Delta t)$ est remplacé par $Q^*(x_{k+1})$ et $P(\Delta x)$ par $P(x_{k+1})$. Ainsi, pour le problème considéré, l'expression (4.22) se met sous la forme

$$Q^*(x_k) = \min_{u_k} [(\lambda^2 x_k^2 + u_k)^2 + \\ + \int_{\Omega(x_k)} Q^*(x_{k+1}) P(x_{k+1} \mid x_k, u_k)\, d\Omega(x_k)], \tag{4.69}$$

où $Q(x_k)$ est la partie de la somme (4.68) de $i = k$ à $i = N$, alors que Q^* est le minimum de Q par rapport à u.

Si l'on constate que $-\infty < x_k < \infty$ et si l'on porte dans (4.69) l'expression de $P(x_{k+1})$ sous la forme (4.67), l'égalité (4.69) devient

$$Q^*(x_k) = \min_{u_k} \Big[(\lambda^2 x_k^2 + u_k^2) + \\ + \int_{-\infty}^{\infty} \frac{Q^*(x_{k+1})}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x_{k+1} - \alpha x_k - u_k)^2}{2\sigma^2}\right\} dx_{k+1}\Big]. \tag{4.70}$$

Connaissant $Q^*(x_N)$ on peut trouver de (4.70) par récurrence les fonctions $Q^*(x_k)$. Puisque (4.68) entraîne

$$Q^*(x_N) = \min_{u_N} \{\lambda^2 x_N^2 + u_N^2\}, \tag{4.71}$$

après avoir calculé la valeur optimale $u_N^* = 0$ et l'avoir portée dans (4.71), on a

$$Q^*(x_N) = \lambda^2 x_N^2. \tag{4.72}$$

Plus généralement, soit

$$Q^*(x_{k+1}) = A_{k+1} + B_{k+1} x_{k+1}^2, \tag{4.73}$$

(4.70) donne alors

$$Q^*(x_k) = \min_{u_k} \Big[(\lambda^2 x_k^2 + u_k^2) + \\ + \int_{-\infty}^{\infty} \frac{A_{k+1} + B_{k+1} x_{k+1}^2}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x_{k+1} - \alpha x_k - u_k)^2}{2\sigma^2}\right\} dx_{k+1} \Big]. \tag{4.74}$$

Posons

$$\left.\begin{aligned} x_{k+1} - \alpha x_k - u_k &= w, \\ dx_{k+1} &= dw. \end{aligned}\right\} \tag{4.75}$$

L'intégrale de (4.74) se transforme alors comme suit :

$$\int_{-\infty}^{\infty} \frac{A_{k+1} + B_{k+1}(w + \alpha x_k + u_k)^2}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw = \\ = \frac{A_{k+1} + B_{k+1}(\alpha x_k + u_k)^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw + \\ + \frac{B_{k+1}}{\sigma\sqrt{2\pi}} 2(\alpha x_k + u_k) \int_{-\infty}^{\infty} w \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw + \\ + \frac{B_{k+1}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} w^2 \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw. \tag{4.76}$$

Substituons les valeurs des intégrales

$$\left.\begin{aligned} &\int_{-\infty}^{\infty} \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw = \sigma\sqrt{2\pi}, \\ &\int_{-\infty}^{\infty} w \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw = 0, \\ &\int_{-\infty}^{\infty} w^2 \exp\left\{-\frac{w^2}{2\sigma^2}\right\} dw = \sigma^3\sqrt{2\pi}. \end{aligned}\right\} \tag{4.77}$$

L'intégrale (4.76) devient égale à

$$A_{k+1} + B_{k+1}(\alpha x_k + u_k)^2 + \sigma^2 B_{k+1}, \tag{4.78}$$

et toute l'expression (4.74) peut s'écrire

$$Q^*(x_k) = \min_{u_k} [\lambda^2 x_k^2 + u_k^2 + A_{k+1} + B_{k+1}(\alpha x_k + u_k)^2 + \sigma^2 B_{k+1}] =$$

$$= \min_{u_k} [(\lambda^2 + B_{k+1}\alpha^2) x_k^2 + A_{k+1} + \sigma^2 B_{k+1} +$$

$$+ u_k^2 (1 + B_{k+1}) + 2B_{k+1}\alpha x_k u_k]. \tag{4.79}$$

Si $(1 + B_{k+1}) > 0$, l'expression (4.79) a un minimum qu'on peut obtenir en égalant à zéro la dérivée $dQ^*(x_k)/du_k$

$$2u_k(1 + B_{k+1}) + 2B_{k+1}\alpha x_k = 0. \tag{4.80}$$

Ainsi la commande optimale

$$u_k^* = -\frac{\alpha B_{k+1} x_k}{1 + B_{k+1}}. \tag{4.81}$$

En portant (4.81) dans (4.79), on a

$$Q^*(x_k) = \left[\lambda^2 + \alpha^2 B_{k+1} - \frac{\alpha^2 B_{k+1}^2}{1 + B_{k+1}}\right] x_k^2 +$$

$$+ A_{k+1} + \sigma^2 B_{k+1} = A_k + B_k x_k^2, \tag{4.82}$$

où

$$\left.\begin{aligned} A_k &= A_{k+1} + \sigma^2 B_{k+1}, \\ B_k &= \lambda^2 + \frac{\alpha^2 B_{k+1}}{1 + B_{k+1}}. \end{aligned}\right\} \tag{4.83}$$

On tire de (4.72)

$$A_N = 0, \qquad B_N = \lambda^2 > 0. \tag{4.84}$$

On en déduit que tous les $B_k > 0$ et donc tous les $A_k > 0$. Ainsi, la commande optimale est définie par (4.81) et les B_k par (4.83) et (4.84). Il est curieux de noter que la linéarité du système et la forme quadratique du critère font que le bruit ξ n'influe pas du tout sur le choix de la commande optimale. En effet, les grandeurs B_k depuis $B_N = \lambda^2$, de même que les valeurs de u_k^*, ne dépendent pas de σ. L'existence du bruit n'intervient que dans la grandeur $Q^*(x_k)$ qui dépend de σ par l'intermédiaire de A_k.

Considérons la classe des systèmes strictement discrets et les problèmes relatifs à leur commande optimale. Ces problèmes ont fait l'objet des ouvrages [4.6-4.8, 4.12]. Il est surtout intéressant d'étudier la commande optimale dans le cas des systèmes discrets markoviens. La théorie et les schémas fonctionnels de tels systèmes sont examinés en détail dans la littérature mathématique et techni-

que théorique (cf. [4.9-4.11]). Les systèmes discrets markoviens de même que les systèmes discrets réguliers décrits dans le chapitre III sont caractérisés par un ensemble des états possibles $q_1, q_2, \ldots$ $\ldots, q_{N+1}$. Désignons par s_t l'état courant à l'instant t. Ce peut être l'un quelconque des états possibles q_j ($j = 1, 2, \ldots, N+1$). L'état s_t change en général avec le temps ($t = 0, 1, 2, \ldots$).

A la différence des systèmes examinés au chapitre III, le processus de transition du système d'un état à l'autre est aléatoire. Cela signifie que, connaissant son état s_t à l'instant t et la commande u_t à ce même instant, nous ne pouvons faire qu'une prévision probabiliste de son état suivant s_{t+1}. Toutes les grandeurs étant discrètes, posons que u_t ne peut prendre qu'une des valeurs possibles α_k ($k = 1, 2, \ldots, m$). Si l'on connaît $s_t = q_j$ et $u_t = \alpha_k$, on ne peut définir que la probabilité de transition de l'objet à l'instant suivant à un certain état q_j, c'est-à-dire la probabilité du fait que s_{t+1} serait égal à q_j. Désignons cette probabilité de transition en fonction de q_j et α_k par

$$p_{ij}(k) = p\{s_{t+1} = q_j \mid s_t = q_i,\ u_t = \alpha_k\}. \tag{4.85}$$

Ainsi $p_{ij}(k)$ est la probabilité du fait que l'application de la commande α_k entraîne la transition du système de l' i-ième état q_i au j-ième état q_j.

L'expression (4.85) impose la remarque suivante. Premièrement, la probabilité de transition p_{ij} est une probabilité conditionnelle qui dépend de l'état s_t à l'instant t et de la commande u_t. La relation entre p_{ij} et s_t détermine le processus aléatoire de variation d'états $s_0, s_1, \ldots, s_t, s_{t+1}, \ldots$ comme un processus markovien. Ensuite, la formule (4.85) n'est pas la plus générale même pour les processus markoviens. Dans le cas général la grandeur p_{ij} peut dépendre également de l'instant t. Bornons-nous pour le moment au cas où p_{ij} ne dépend pas explicitement de t. Enfin les expressions (4.85) pour différents i et j ne peuvent pas être quelconques. D'abord les p_{ij} reposent entre zéro et l'unité: $0 \leqslant p_{ij} \leqslant 1$. D'autre part le système passe inévitablement de l'état s_t à un autre état quelconque s_{t+1}. Dans un cas particulier, ce nouvel état peut coïncider avec l'ancien, mais le système passe obligatoirement à un état quelconque. C'est pourquoi la somme des probabilités p_{ij} pour i considéré par rapport à tous les j doit être égale à l'unité

$$\sum_{j=1}^{N+1} p_{ij} = 1. \tag{4.86}$$

On dit que l'on dispose de l'information maximale sur un système discret markovien si l'on connaît à l'instant t son état s_t et la commande d'entrée u_t. Cette information est incomplète car, elle ne définit pas d'une façon univoque le comportement futur du système

mais permet d'obtenir seulement ses caractéristiques probabilistes. Or, en principe il est impossible de recevoir des renseignements plus précis sur ce système.

Le système discret markovien est décrit parfaitement par un certain nombre de matrices (m au total)

$$\overline{\overline{P}}_k = \| p_{ij}(k) \| \tag{4.87}$$

des probabilités de transition pour différentes actions de commande α_k ($k = 1, \ldots, m$). Chacune des matrices $\overline{\overline{P}}_k$ peut être remplacée par un diagramme sur lequel les états possibles q_j ($j = 1, \ldots, N+1$) sont marqués par des cercles affectés d'indices j. Par exemple, sur le diagramme de la figure 4.6 sont montrés deux états possibles 1 et 2. Les transitions possibles d'un état à l'autre sont indiquées

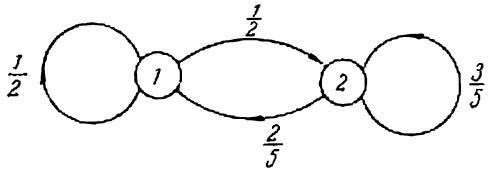


Fig. 4.6.

par les flèches et les probabilités de ces transitions figurent près de ces flèches. Par exemple, la transition de l'état 2 à l'état 1 peut avoir lieu avec une probabilité $\frac{2}{5}$ et à l'état 2 avec la probabilité $\frac{3}{5}$, c'est-à-dire que le système garde le même état avec la probabilité $\frac{3}{5}$.

Outre les probabilités de transition, il est également intéressant d'étudier les probabilités d'états d'un système markovien. Désignons par $p_i(t)$ la probabilité du fait que l'état s_t à l'instant t est q_i, c'est-à-dire l' i-ème des états possibles. Puisqu'à l'instant t correspond obligatoirement un état quelconque, la somme des probabilités $p_i(t)$ est égale à l'unité

$$\sum_{j=1}^{N+1} p_i(t) = 1. \tag{4.88}$$

Pour définir un processus markovien il faut imposer encore les probabilités d'états initiales. Si on se donne notamment un certain état initial, q_j par exemple, cela signifie que $p_j(0) = 1$, tandis que toutes les autres $p_i(0)$ sont nulles ($i \neq j$). Mais dès l'instant suivant la transition de l'état q_j à un autre état est déjà possible. C'est pourquoi avec $t > 0$ la probabilité pour le système de se trouver dans un état déterminé est « étalée » sur différents états possibles.

Il n'est pas difficile de déduire la formule qui permet de calculer toutes les $p_j(t+1)$ connaissant les $p_i(t)$ ($i = 1, \ldots, N+1$)

à l'instant t et sachant que $u = \alpha_k$. Considérons l'état q_j à l'instant $(t + 1)$. Les probabilités $p_i(t)$ étant connues, cherchons la probabilité de l'événement composé qui consiste à passer de l'état q_i à l'instant t à l'état q_j à l'instant $(t + 1)$. La probabilité concernée de l'événement composé est égale au produit de la probabilité inconditionnelle pour le système de se trouver à l'instant t en état q_i, c'est-à-dire de $p_i(t)$, par la probabilité conditionnelle $p_{ij}(k)$ de transition de l' i-ième état au j-ième état. Ainsi la probabilité de cet événement composé est égale à $p_i(t) \cdot p_{ij}(k)$. Comme la probabilité $p_j(t + 1)$ de se trouver en état q_j à l'instant $(t + 1)$ est la somme des probabilités de transition à l état q_j à partir de tous les états q_i $(i = 1, \ldots, N + 1)$, les $N + 1$ égalités

$$p_j(t+1) = \sum_{i=1}^{N+1} p_i(t)\, p_{ij}(k) \quad (j = 1, \ldots, N+1, \quad t = 0, 1, \ldots) \tag{4.89}$$

sont justifiées. Ces expressions permettent d'obtenir, à partir des probabilités d'états initiales $p_i(0)$, les probabilités d'états $p_i(1)$ pour $t=1$. Ensuite, d'après les mêmes formules (mais, au fond, pour une autre commande α_k), on peut calculer les grandeurs $p_i(2)$, etc.

Introduisons le vecteur des probabilités d'états (c'est un vecteur ligne)

$$\overline{p}(t) = (p_1(t), p_2(t), \ldots, p_{N+1}(t)). \tag{4.90}$$

Les égalités (4.89) se remplacent par une seule égalité vectorielle

$$\overline{p}(t + 1) = \overline{p}(t)\overline{\overline{P}}_k, \tag{4.91}$$

où $\overline{\overline{P}}_k$ est la matrice définie par l'expression (4.87). Examinons à titre d'exemple le cas où $u = \alpha_k$ est invariable et indépendant de t. On a

$$\left.\begin{aligned} \overline{p}(1) &= \overline{p}(0)\overline{\overline{P}}_k, \\ \overline{p}(2) &= \overline{p}(1)\overline{\overline{P}}_k = \overline{p}(0)\overline{\overline{P}}_k^2 \end{aligned}\right\} \tag{4.92}$$

et, en général,

$$\overline{p}(t) = \overline{p}(0)\overline{\overline{P}}_k^t, \tag{4.93}$$

où la matrice $\overline{\overline{P}}_k^t$ est la t-ième puissance de la matrice $\overline{\overline{P}}_k$. L'équation (4.93) donne en principe $\overline{p}(t)$ pour tout t. Soit, par exemple, pour le schéma représenté sur la figure 4.6

$$p_1(0) = 1, \quad p_2(0) = 0. \tag{4.94}$$

Cela signifie qu'à l'instant initial $t = 0$ le système est en état 1. Le vecteur des probabilités d'états à l'instant initial

$$\overline{p}(0) = (1, 0). \tag{4.95}$$

La matrice $\overline{\overline{P}}_k$ s'écrit (cf. fig. 4.6)

$$\overline{\overline{P}}_k = \begin{vmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{3}{5} \end{vmatrix}. \tag{4.96}$$

Ensuite, d'après les règles connues de la multiplication des matrices on trouve

$$\overline{p}(1) = \overline{p}(0)\,\overline{\overline{P}}_k = \left| 1,\ 0 \right| \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{3}{5} \end{vmatrix} = \left| \frac{1}{2},\ \frac{1}{2} \right|, \tag{4.97}$$

$$\overline{p}(2) = \overline{p}(1)\,\overline{\overline{P}}_k = \left| \frac{1}{2},\ \frac{1}{2} \right| \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{5} & \frac{3}{5} \end{vmatrix} = \left| \frac{9}{20},\ \frac{11}{20} \right|, \tag{4.98}$$

etc. Le vecteur doit être considéré ici comme une matrice ligne. Les valeurs $p_1(t)$ et $p_2(t)$ sont consignées sur la table suivante

t	0	1	2	3	4	5
$p_1(t)$	1	0,5	0,45	0,445	0,4445	0,44445
$p_2(t)$	0	0,5	0,55	0,555	0,5555	0,55555

Cette table montre qu'à mesure que t croît, les grandeurs $p_1(t)$ et $p_2(t)$ tendent vers des limites constantes: $p_1(t) \to 0{,}4444\ldots$, et $p_2(t) \to 0{,}5555\ldots$ Ces valeurs constituent ce qu'on appelle les probabilités limites. On peut montrer que les valeurs des probabilités limites sont les mêmes dans le système considéré, quelles que soient les conditions initiales $p_1(0)$ et $p_2(0)$. Du point de vue physique les probabilités limites sont les probabilités dans un régime stationnaire. Le processus markovien dont les probabilités limites ne dépendent pas des conditions initiales est dit processus totalement ergodique [4.7]. On en déduit que le processus représenté par le diagramme de la figure 4.6 est totalement ergodique.

Pour établir les probabilités limites d'un processus totalement ergodique on n'a pas besoin de calculer les suites $p_j(t)$ ($t = 0, 1, 2, \ldots$). Considérons, par exemple, les équations (4.91) relatives au schéma de la figure 4.6. Ecrivons-les sous la forme de deux équations du type (4.89)

$$\left.\begin{aligned} p_1(t+1) &= \frac{1}{2}\,p_1(t) + \frac{2}{5}\,p_2(t), \\ p_2(t+1) &= \frac{1}{2}\,p_1(t) + \frac{3}{5}\,p_2(t). \end{aligned}\right\} \tag{4.99}$$

Lorsque le processus transitoire s'achève, les probabilités p_1 et p_2 prennent les valeurs permanentes p_{1f} et p_{2f}. En remplaçant dans (4.99) $p_1(t+1)$ et $p_1(t)$ par p_{1f} et $p_2(t+1)$, $p_2(t)$ par p_{2f}, on obtient les équations

$$\left.\begin{aligned} p_{1f} &= \frac{1}{2}p_{1f} + \frac{2}{5}p_{2f}, \\ p_{2f} &= \frac{1}{2}p_{1f} + \frac{3}{5}p_{2f}. \end{aligned}\right\} \tag{4.100}$$

La résolution de ces équations donne

$$p_{1f} = \frac{4}{9} = 0{,}4444\ldots, \quad p_{2f} = \frac{5}{9} = 0{,}5555\ldots, \tag{4.101}$$

ce qui coïncide avec les limites vers lesquelles tendent pour $t \to \infty$ les valeurs de $p_j(t)$ indiquées dans la table.

Dans cet exemple les deux probabilités limites p_{1f} et p_{2f} sont différentes du zéro. Mais des cas peuvent se présenter lorsque l'une

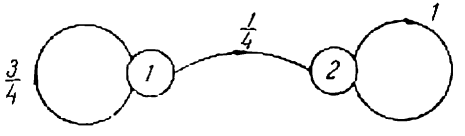


Fig. 4.7.

d'elles est égale à un et les autres sont nulles. Cela signifie que le régime stationnaire du système s'établit dans un état défini. Tel est, par exemple, le système dont le diagramme est représenté sur la figure 4.7. Il est clair sans aucun calcul que, parvenu à l'état *2* (la probabilité de cet état est supérieure à zéro), le système n'en sortira jamais plus.

Des cas sont possibles lorsque la répartition des probabilités limites dépend des conditions initiales. Tel est, par exemple, le système du diagramme représenté sur la figure 4.8. Si à l'instant initial le système est en état *1*, il y reste; mais si à l'instant initial il est en état *2*, il y reste également.

Un régime stationnaire peut donner lieu à des cycles limites. Considérons, par exemple, le diagramme de la figure 4.9. Quelles que soient les conditions initiales, le système sera en définitive « entraîné » dans le cycle limite composé d'alternances des états *2* et *3*.

On voit de ce qui vient d'être dit que la définition des $p_j(t)$ se ramène à l'étude des équations aux différences finies du type (4.89). Ces équations ont été étudiées dans les ouvrages traitant des processus markoviens [4.13, 4.14], dans la théorie des équations

aux différences finies [4.15], dans la théorie des systèmes impulsionnels [4.11], ainsi que dans la théorie d'apprentissage en physiologie [4.17, 4.18].

Examinons la méthode générale de recherche de la commande optimale pour un système discret markovien [4.12]. Supposons

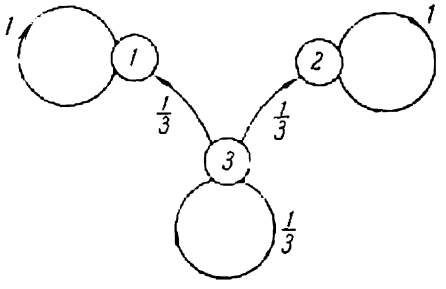


Fig. 4.8.

que le système possède $N+1$ états $q_1, q_2, \ldots, q_{N+1}$ dont q_{N+1} est l'état final imposé.

Associons à une transition quelconque de q_i à q_j, pour laquelle la probabilité $p_{ij}(k)$ est positive, un certain « coût » positif $c_{ij}(k) > 0$. De plus, le coût associé aux transitions de l'état final à ce

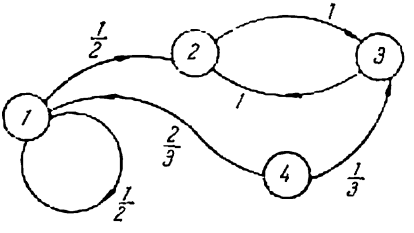


Fig. 4.9.

même état est considéré nul ; d'autre part on adopte encore que les transitions de l'état final à d'autres états sont inexistantes. Il vient

$$p_{N+1,\,N+1}(k) = 1 \quad (k = 1, 2, \ldots, m) \tag{4.102}$$

pour toute commande $u_k = \alpha_k$.

Dans le cas particulier, pour tout i, j, k tel que $p_{ij}(k) \neq 0$, $c_{ij}(k) = 1$. Le coût total du processus est alors d'autant plus grand qu'il y a plus de transitions, c'est-à-dire que la durée du processus est plus grande.

A l'instant $t = 0$, le système est en un certain état initial s_0, mesurable comme tous les s_t $(t \geqslant 0)$. Une information précise sur

l'état courant s_t est canalisée vers l'organe de commande qui élabore la consigne u_t. Toutes les $p_{ij}(k)$ étant connues, la connaissance de l'état courant permet de calculer la probabilité de transition à tous les autres états. Ainsi l'information disponible à l'instant courant sur le système gouverné est maximale possible.

Pendant l pas successifs on observe une certaine suite des états du système $(s_0, s_1, \ldots, s_l)$, ainsi qu'une suite des commandes $(u_0, u_1, \ldots, u_{l-1})$. Appelons possible la suite des commandes et des états $(\alpha_{k0}, \alpha_{k1}, \ldots, \alpha_{k,l-1}, q_{j0}, q_{j1}, \ldots, q_{jl})$ si elle est réalisable, c'est-à-dire si la probabilité conditionnelle de l'apparition de cette suite des commandes $(\alpha_{k0}, \alpha_{k1}, \ldots, \alpha_{k,l-1})$ et des états $(q_{j0}, q_{j1}, \ldots, q_{jl})$, pour l'état initial q_{j0}, est supérieure à zéro.

A chaque suite possible de longueur l nous associons le coût c de la manière suivante :

$$c = \sum_{\nu=0}^{l} c_{i_\nu i_{\nu+1}}(k), \tag{4.103}$$

$c_{i_\nu i_{\nu+1}}$ étant ici le coût d'une transition de $s_\nu = q_{i_\nu}$ à $s_{\nu+1} = q_{i_{\nu+1}}$, lorsque la commande à l'entrée du système est $u_\nu = \alpha_{k_\nu}$. Lorsque ce dernier atteint l'état final, le coût c ne croît plus. En effet, les coûts des transitions ultérieures de l'état q_{N+1} à ce même état sont nuls. La formule (4.103) permet de définir également la valeur de c pour les processus dans lesquels l'état final n'est pas atteint. Dans ces conditions, si le nombre de pas est infini, le coût total du processus devient infiniment grand.

Dans [4.12] on appelle stratégie, ou politique, l'algorithme de commande, c'est-à-dire la loi

$$u_t = \pi(s_t), \tag{4.104}$$

qui associe une certaine commande à chaque état du système, sauf l'état final. Cette commande u_t doit être donnée si le système se trouve en état s_t. La loi (4.104) est une analogie parfaite de l'algorithme

$$\bar{u} = \bar{u}(\bar{x}) \tag{4.105}$$

relatif à un système continu. En effet, à chaque point $\bar{x}$ de l'espace de phase l'algorithme (4.105) fait correspondre une commande définie $\bar{u}$. Pourtant, pour le système discret décrit plus haut, l'espace de phase est également discret et se transforme en un ensemble fini des états *1*, *2*, . . . (cf. fig. 4.6 à 4.9). Les deux lois (4.104) et (4.105) sont caractéristiques des systèmes à information complète ou maximale sur l'objet, lorsque la connaissance de la « préhistoire », c'est-à-dire de la suite des états $s_0, \ldots, s_{t-1}$ n'ajoute rien à l'information disponible sur l'objet. Pour notre part nous n'allons pas appeler stratégie les lois (4.104) et (4.105).

Il est commode de considérer l'ensemble des expressions π_1, π_2, . . ., π_N du type (4.104) pour les états possibles q_1, q_2, . . ., q_N du système comme le vecteur

$$\bar{\pi} = (\pi_1, \pi_2, \ldots, \pi_N). \tag{4.106}$$

On pose ici que la commande π_i est appliquée lorsque le système est en état q_i. De plus, dans ce qui suit nous employons les notations $p_{ij}(\pi_i)$ et $c_{ij}(\pi_i)$ au lieu de $p_{ij}(k)$ et $c_{ij}(k)$. Nous avons adopté encore que la commande prend l'une de ses valeurs possibles α_1, ..., α_m.

Les transitions d'un état à l'autre étant aléatoires, le coût total c du processus est également une grandeur aléatoire dont la répartition dépend du type de l'algorithme $\bar{\pi}$. Adoptons comme critère d'optimalité l'espérance mathématique du coût de l'état final $X_i(\bar{\pi})$. Cette grandeur est fonction de l'état initial q_i et de l'algorithme adopté $\bar{\pi}$. Désignons par $\bar{X}(\bar{\pi})$ le vecteur

$$\bar{X}(\bar{\pi}) = (X_1(\bar{\pi}), X_2(\bar{\pi}), \ldots, X_N(\bar{\pi})), \tag{4.107}$$

dont les coordonnées sont les espérances mathématiques du coût au départ à partir des états q_1, q_2, . . ., q_N. Pour abréger l'écriture appelons $\bar{X}(\bar{\pi})$ vecteur de coût. Soit $\bar{X}$ et $\bar{Y}$ deux vecteurs de coût associés à deux algorithmes différents. Alors l'expression $\bar{X} \leqslant \bar{Y}$ signifie que $X_i \leqslant Y_i$ pour tout $i = 1, 2, \ldots, N$. Ainsi l'algorithme associé au vecteur $\bar{X}$ est uniformément meilleur que celui associé au vecteur $\bar{Y}$. Si nous considérons le coût relatif au processus de longueur fixée l, nous pouvons le noter $X_i(\pi, l)$ ou plus simplement $X_i(l)$.

Nous avons dit que l'algorithme $\bar{\pi}^0$ est optimal s'il vérifie la condition

$$\bar{X}(\bar{\pi}^0) \leqslant \bar{X}(\bar{\pi}) \tag{4.108}$$

pour tout $\bar{\pi}$. Pour l'algorithme optimal $\bar{\pi}^0$ on peut déduire une relation du même type que l'équation de Bellman. Examinons d'abord un algorithme quelconque $\bar{\pi}$ et le coût éventuel $X_i(\bar{\pi})$ lors du démarrage à partir de l'état initial q_i. La probabilité de passer pendant le premier pas de q_i à un certain état q_j est $p_{ij}(\pi_i)$, et le coût de cette transition est $c_{ij}(\pi_i)$. Supposons d'abord que le premier pas a conduit le système à l'état q_j. Le coût total du processus est donné alors par la somme

$$c_{ij}(\pi_i) + X_j(\bar{\pi}), \tag{4.109}$$

où $X_j(\bar{\pi})$ est l'espérance mathématique du coût du processus qui part de l'état q_j. La grandeur définie par l'équation (4.109) est aléatoire du fait qu'après le premier pas le système peut se trouver

dans différents états q_j. En prenant la moyenne de la grandeur (4.109) on obtient pour l'espérance mathématique $X_i(\overline{\pi})$ l'équation suivanvante :

$$X_i(\overline{\pi}) = M\{c_{ij}(\pi_i) + X_j(\overline{\pi})\} \tag{4.110}$$

ou sous une forme développée

$$X_i(\overline{\pi}) = \sum_{j=1}^{N+1} p_{ij}(\pi_i)[c_{ij}(\pi_i) + X_j(\overline{\pi})] \quad (i \quad 1, \ldots, N). \tag{4.111}$$

Etant donné que $X_{N+1}(\overline{\pi}) = 0$, on a

$$X_i(\overline{\pi}) = \sum_{j=1}^{N} p_{ij}(\pi_i) X_j(\overline{\pi}) + \sum_{j=1}^{N+1} p_{ij}(\pi_i) c_{ij}(\pi_i) \quad (i = 1, \ldots, N). \tag{4.112}$$

Les N égalités scalaires peuvent être remplacées par une égalité vectorielle

$$\overline{X}(\overline{\pi}) = \overline{\overline{P}}(\overline{\pi})\,\overline{X}(\overline{\pi}) + \overline{c}(\overline{\pi}), \tag{4.113}$$

où $\overline{\overline{P}}(\overline{\pi})$ est la matrice carrée à N lignes d'éléments $p_{ij}(\pi_i)$ et $\overline{c}(\overline{\pi})$ le vecteur dont l'i-ième coordonnée s'écrit

$$\sum_{j=1}^{N+1} p_{ij}(\pi_i) \cdot c_{ij}(\pi_i) = c_i(\overline{\pi}). \tag{4.114}$$

La grandeur $c_i(\overline{\pi})$ est l'espérance mathématique du coût du premier pas lors du départ de q_i et en utilisant l'algorithme $\overline{\pi}$. On peut alors appeler $\overline{c}(\overline{\pi})$ vecteur de coût du premier pas.

L'égalité (4.113) est valable pour n'importe quel algorithme. Désignons par $\overline{X}^0(\overline{\pi}^0)$ le vecteur de coût relatif à l'algorithme optimal $\overline{\pi}^0$.

Etablissons la relation pour $\overline{X}^0(\overline{\pi}^0) = \overline{X}^0$ en raisonnant de la manière suivante. Supposons que le premier pas à partir de q_i, effectué en appliquant la commande π_i, soit fixé. Le système atteint l'état q_j, alors que les pas suivants se font d'après une stratégie optimale. D'une façon analogue à (4.109), l'espérance mathématique du coût du processus s'écrit

$$c_{ij}(\pi_i) + X_j^0(\overline{\pi}^0). \tag{4.115}$$

L'espérance mathématique de cette grandeur, égale à X_i, s'obtient sous la forme

$$X_i = M\{c_{ij}(\pi_i) + X_j^0\} = \sum_{j=1}^{N+1} p_{ij}[c_{ij}(\pi_i) + X_j^0]. \tag{4.116}$$

Pour obtenir l'algorithme optimal X_i^0 il faut minimiser le second membre de cette égalité par rapport à toutes les valeurs possibles de π_i. Ainsi

$$X_i^0 = \min_{\pi_i} \{ \sum_{j=1}^{N+1} p_{ij} [c_{ij}(\pi_i) + X_j^0] \}. \tag{4.117}$$

Certes, le minimum peut être pris par rapport à $\bar{\pi}$ et non pas à π_i sans que l'égalité (4.117) change. Elle peut donc s'écrire sous la forme vectorielle suivante :

$$\bar{X}^0 = \min_{\bar{\pi}} [\bar{\bar{P}}(\bar{\pi}) \bar{X}^0 + \bar{c}(\bar{\pi})]. \tag{4.118}$$

En résolvant cette équation on obtient $\bar{X}^0$, tout en définissant l'algorithme optimal $\bar{\pi}^0$. Pour rechercher la solution on peut mettre à profit le fait que $\bar{X}^0$, comme on le voit de (4.118), est un point fixe de la transformation

$$T(\bar{X}) = \min_{\bar{\pi}} [\bar{\bar{P}}(\bar{\pi}) \bar{X} + \bar{c}(\bar{\pi})]. \tag{4.119}$$

S'étant imposé une approximation nulle de $\bar{X}$ et effectuant les opérations écrites dans le second membre de cette dernière égalité, on obtient la première approximation $T(\bar{X})$ du vecteur $\bar{X}^0$. En remplaçant $\bar{X}$ dans le second membre de (4.119) par cette première approximation et après avoir effectué les opérations indiquées dans (4.119), on obtient la deuxième approximation $T^2(\bar{X})$, etc. On peut supposer ainsi que la méthode itérative permet de s'approcher aussi près que l'on veut de la grandeur imposée $\bar{X}^0$. La minimisation du second membre de (4.119) permet de déterminer en même temps l'algorithme $\bar{\pi}$, qui s'approche de plus en plus de l'optimal. Ainsi

$$\bar{X}^0 = \lim_{r \to \infty} T^r(\bar{X}). \tag{4.120}$$

Bien entendu, les hypothèses énoncées ne peuvent d'aucune façon remplacer la démonstration. Mais existe-t-il en général pour ce problème un algorithme uniformément optimal? Et s'il existe, est-il l'algorithme unique? Le processus itératif exposé plus haut converge-t-il vers cet algorithme indépendamment de la valeur initiale de $\bar{X}$?

Dans [4.12] on trouve une réponse exhaustive à toutes ces questions. Il se trouve que dans certaines conditions un tel algorithme existe bien et il est unique. Le processus itératif converge quelle que soit la valeur initiale de $\bar{X}$.

Le résultat principal obtenu dans [4.12] et dont nous omettons la démonstration peut être énoncé de la manière suivante.

Appelons algorithme uniformément optimal l'algorithme pour lequel, à partir d'un état initial quelconque, le système peut atteindre l'état final avec une probabilité positive.

S'il existe au moins un algorithme uniformément optimal, l'équation (4.118) a une solution unique définie par la formule (4.120), où $\overline{X}$ est le vecteur initial et T est donné par l'expression (4.119).

L'intérêt de cette dernière expression est grand car elle fournit un moyen efficace de calcul de $\overline{X}^0$ et de l'algorithme $\overline{\pi}^0$ qui lui est associé.

Les raisonnements exposés peuvent être généralisés au cas où l'état final du système varie en fonction du temps. Nous obtenons alors le problème de « coup au but » dans le cas d'une cible animée d'un mouvement. Supposons que la consigne transmise à l'organe de commande à l'instant t soit une grandeur discrète ρ_t appartenant à l'ensemble $q_1, \ldots, q_N$. (L'état q_{N+1} n'existe déjà plus. Au lieu de q_{N+1}, à tout instant figure en tant qu'état final l'un des états q_i.) Cela signifie que ρ_t est la valeur courante de la consigne. Elle peut varier d'une façon aléatoire et constitue un processus markovien à probabilité transitoire $p(\rho_{t+1} \mid \rho_t)$. C'est une probabilité conditionnelle du fait qu'à l'instant $(t+1)$ l'état final du système est ρ_{t+1}, si à l'instant t il était ρ_t. Pour généraliser à ce cas la méthode exposée ci-dessus définissons un certain état « vectoriel » complexe $\overline{s}_t = (s_t, \rho_t)$, c'est-à-dire un vecteur de coordonnées s_t et ρ_t.

La probabilité conditionnelle d'un tel vecteur est donnée par l'expression qui se déduit du théorème du produit des probabilités :

$$\begin{aligned} p(\overline{s}_{t+1} \mid \overline{s}_t, u_t) &= p(s_{t+1}, \rho_{t+1} \mid s_t, \rho_t, u_t) = \\ &= p(s_{t+1} \mid s_t, \rho_t, u_t) \cdot p(\rho_{t+1} \mid s_{t+1}, s_t, \rho_t, u_t) = \\ &= p(s_{t+1} \mid s_t, u_t) \cdot p(\rho_{t+1} \mid \rho_t). \end{aligned} \tag{4.121}$$

Comme la probabilité de s_{t+1} ne dépend que de s_t et de u_t et la probabilité de ρ_{t+1} que de ρ_t, les coordonnées s_{t+1} et ρ_{t+1} sont indépendantes.

Le vecteur $\overline{s}_t$ peut avoir une des valeurs possibles $\overline{q}_\lambda = (q_i, q_j)$. Il y a au total N^2 états. Par conséquent, le vecteur $\overline{s}_t$ est un processus markovien à N^2 états et la probabilité de transition définie par la formule (4.121) est considérée comme connue.

Lorsque l'état s_t du système initial coïncide avec la consigne ρ_t, le processus markovien $\overline{s}_t$ donne lieu à l'un des états $\overline{s}_t = (q_i, q_i)$. Le nombre de ces états est N ; chacun d'entre eux correspond au «coup au but ». L'ensemble de tous ces états (q_i, q_i) $(i = 1, \ldots, N)$ constitue un domaine que nous allons nommer Q. Par conséquent, le problème de « coup au but » optimal se ramène au problème de transition optimale du système $\overline{s}_t$ de l'état initial imposé à l'un des états quelconque du domaine Q. Certes, pour le nouveau système

on peut reprendre la notation s_t au lieu de $\bar{s}_t$. Le problème reste au fond le même à cette différence près qu'au lieu de rejoindre un état final, le système atteint le domaine imposé Q constitué de plusieurs états.

Soient les états appartenant au domaine Q notés $q_{L+1}, \ldots, q_N$. Le processus s'achevant lorsque l'état s_t du système appartient pour la première fois à Q, les états $q_{L+1}, \ldots, q_N$ peuvent s'appeler *points d'absorption*. En effet, le processus ne quitte plus le domaine Q. Le coût de transition d'un point quelconque du domaine Q à un autre point de ce domaine est évidemment égal à zéro :

$$X_{L+1} = X_{L+2} = \ldots = X_N = 0. \tag{4.122}$$

L'équation des coûts $X_i(\bar{\pi})$ $(i = 1, \ldots, L)$ des trajectoires menant dans le domaine Q, qui se déduit d'une manière analogue à la relation (4.112), peut s'écrire :

$$X_i(\bar{\pi}) = \sum_{j=1}^{L} p_{ij}(\pi_i) X_j(\bar{\pi}) + \sum_{j=1}^{L} p_{ij}(\pi_i) c_{ij} \qquad (i = 1, 2, \ldots, L). \tag{4.123}$$

Ces équations s'obtiennent à partir de (4.112) si l'on tient compte de (4.122) et si l'on pose $p_{ij}(\pi_i) = 0$ pour $i = L + 1, \ldots, N$.

Sous la forme vectorielle la relation (4.123) s'écrit

$$\bar{X}(\bar{\pi}) = \bar{\bar{P}}^*(\bar{\pi}) \bar{X}(\bar{\pi}) + \bar{c}^*(\bar{\pi}), \tag{4.124}$$

où

$$\left.\begin{aligned} \bar{\pi} &= (\pi_1, \ldots, \pi_L), \\ \bar{X}(\bar{\pi}) &= (X_1(\bar{\pi}), \ldots, X_L(\bar{\pi})), \\ \bar{c}^*(\bar{\pi}) &= (c_1^*(\bar{\pi}), \ldots, c_L^*(\bar{\pi})), \end{aligned}\right\} \tag{4.125}$$

et $\bar{\bar{P}}^*(\pi)$ est la matrice d'éléments $p_{ij}(\pi_i)$ $(i, j = 1, \ldots, L)$. Enfin

$$c_i^*(\pi_i) = \sum_{j=1}^{L} p_{ij}(\pi_i) c_{ij}(\pi_i) \qquad (i = 1, 2, \ldots, L). \tag{4.126}$$

Dans ce cas-là également le vecteur du minimum du coût est défini par une équation analogue à (4.118) :

$$\bar{X}^0 = \min_{\bar{\pi}} [\bar{\bar{P}}^*(\bar{\pi}) \bar{X}^0 + \bar{c}^*(\bar{\pi})]. \tag{4.127}$$

Dans le cas considéré on remplace d'abord le système initial par un système plus complexe et puis on recherche la solution du problème pour le système obtenu.

Dans le cas traité plus haut nous avons examiné un processus transitoire dan sun système discret markovien et défini la commande

optimale minimisant un certain critère (le temps, par exemple) associé au processus transitoire. Toutefois, on peut résoudre non seulement les problèmes relatifs au processus transitoire, mais encore les problèmes concernant un processus stationnaire dans un système discret markovien [4.7]. Soit, par exemple, chaque transition de l'état q_i à l'état q_j apporte non pas une perte c_{ij} mais un gain r_{ij} (le gain peut être considéré comme une perte négative, ou un coût négatif de transition). Au lieu de l'espérance mathématique du coût total X_i du processus partant de l'état q_i on peut introduire l'espérance mathématique du gain total $v_i(n)$ pour n pas.

Considérons n pas (cycles, étapes), l'état initial étant q_i. Supposons d'abord que le premier pas consiste dans la transition de q_i à un état fixé q_j avec un gain r_{ij}. L'espérance mathématique du gain total pour les $n-1$ pas restants étant égale à $v_j(n-1)$, l'espérance mathématique du gain total dans les conditions considérées peut s'obtenir sous la forme

$$r_{ij} + v_j(n-1). \tag{4.128}$$

Pourtant, en réalité, la transition de q_i à q_j est aléatoire. La grandeur définie par (4.128) est donc également aléatoire. En prenant sa moyenne sur l'ensemble des états q_j munis de probabilités p_{ij}, on obtient l'espérance mathématique $v_i(n)$ du gain total pour le démarrage de l'état q_i

$$v_i(n) = M\{r_{ij} + v_j(n-1)\} = \sum_{j=1}^{N} p_{ij}[r_{ij} + v_j(n-1)] =$$

$$= h_i + \sum_{j=1}^{N} p_{ij} v_j(n-1)$$

$$(i = 1, 2, \ldots, N, \ n = 1, 2, 3, \ldots), \tag{4.129}$$

où

$$h_i = \sum_{j=1}^{N} p_{ij} r_{ij} \quad (i = 1, \ldots, N). \tag{4.130}$$

L'égalité (4.129) est analogue à (4.112) et (4.130) à (4.114). Introduisons les vecteurs colonnes $\bar{v}(n)$ et $\bar{h}$

$$\bar{v}(n) = \begin{vmatrix} v_1(n) \\ \vdots \\ v_N(n) \end{vmatrix}, \quad \bar{h} = \begin{vmatrix} h_1 \\ \vdots \\ h_N \end{vmatrix}, \tag{4.131}$$

ainsi que la matrice $\bar{\bar{P}}$ aux éléments p_{ij} comptant N colonnes et N lignes. On a alors l'égalité vectorielle

$$\bar{v}(n) = \bar{h} + \bar{\bar{P}}\bar{v}(n-1), \tag{4.132}$$

qui est analogue à (4.113). Connaissant $\bar{h}$ et $\bar{\bar{P}}$ et sachant que $\bar{v}(1) = \bar{h}$, on tire de l'égalité (4.132) d'abord $\bar{v}(2)$, puis $\bar{v}(3)$, etc. Ce problème ne perd pas de sens pour l'accroissement illimité de n. Pour des n grands on est en présence d'un processus stationnaire.

Maintenant tenons compte du fait que p_{ij} et r_{ij} dépendent de la commande $u_n = \alpha_k$ $(k = 1, \ldots, m)$. Recherchons la commande optimale telle que le gain total $v(n)$ soit maximal. Ecrivons les grandeurs p_{ij} et r_{ij} sous la forme $p_{ij}(k)$ et $r_{ij}(k)$ car elles dépendent de α_k. Raisonnons ensuite de même que lors de la déduction de l'égalité (4.118) mais en remplaçant le minimum par le maximum. Soit le premier pas de q_i à q_j fixé, ainsi que la commande $u_0 = \alpha_k$ et de plus la commande est optimale à tout instant ultérieur. L'espérance mathématique $v(n)$ du gain en n pas par analogie avec (4.128) s'écrit

$$r_{ij}(k) + v_j^0(n-1), \tag{4.133}$$

où $v_j^0(n-1)$ est l'espérance mathématique du gain en $n-1$ pas lors du départ à partir de q_j et sous la condition que la commande soit optimale. En choisissant la commande $u = \alpha_k$ au *premier* pas de façon qu'elle maximise l'espérance mathématique de l'expression (4.133), on aboutit à l'égalité

$$v_i^0(n) = \max_k \sum_{j=1}^{N} p_{ij}(k)\,[r_{ij}(k) + v_j^0(n-1)] \tag{4.134}$$
$$(i = 1, \ldots, N)$$

ou sous forme vectorielle

$$\bar{v}^0(n) = \max_k \{\bar{\bar{P}}_k \bar{v}^0(n-1) + \bar{h}\}. \tag{4.135}$$

En procédant de proche en proche on calcule à l'aide des égalités (4.134) ou (4.135) l'espérance mathématique du gain $\bar{v}^0(n)$. En effet, si l'on connaît $\bar{v}^0(0) = 0$, on peut déterminer successivement $\bar{v}^0(1)$, $\bar{v}^0(2)$, $\bar{v}^0(3)$, etc. La maximisation simultanée par rapport à u_k donne la commande optimale.

Nous avons déjà dit que la grandeur

$$h_i = h_i(k) = \sum_{j=1}^{N} p_{ij}(k)\, r_{ij}(k) \tag{4.136}$$

dépend de la commande $u_n = \alpha_k$ $(k = 1, \ldots, m)$. A son tour, le numéro k de la commande dépend de n. En introduisant la notation $h_i(k)$ on peut écrire (4.134) sous la forme

$$v_i^0(n) = \max_k \left[h_i(k) + \sum_{j=1}^{N} p_{ij}(k)\, v_j^0(n-1)\right]. \tag{4.137}$$

Etant donné que les $p_{ij}(k)$ dépendent de k, les deux termes entre crochets de l'expression (4.137) en dépendent également.

Considérons l'exemple tiré de [4.7]. Soient deux commandes possibles $u = \alpha_k$, $k = 1, 2$, et deux états q_1 et q_2. Supposons que dans le cas $k = 1$ les matrices des $p_{ij}(1)$ et $r_{ij}(1)$ s'écrivent

$$\bar{\bar{P}}_1 = \| p_{ij}(1) \| = \begin{vmatrix} 0{,}5 & 0{,}5 \\ 0{,}4 & 0{,}6 \end{vmatrix}, \quad \bar{\bar{R}}_1 = \| r_{ij}(1) \| = \begin{vmatrix} 9 & 3 \\ 3 & -7 \end{vmatrix}. \tag{4.138}$$

Dans le cas $k = 2$ on a

$$\bar{\bar{P}}_2 = \| p_{ij}(2) \| = \begin{vmatrix} 0{,}8 & 0{,}2 \\ 0{,}7 & 0{,}3 \end{vmatrix}, \quad \bar{\bar{R}}_2 = \| r_{ij}(2) \| = \begin{vmatrix} 4 & 4 \\ 1 & -19 \end{vmatrix}. \tag{4.139}$$

Calculons les valeurs de $h_i(k)$

$$\left.\begin{aligned} h_1(1) &= p_{11}(1)\, r_{11}(1) + p_{12}(1)\, r_{12}(1) = 0{,}5 \cdot 9 + 0{,}5 \cdot 3 = 6, \\ h_2(1) &= p_{21}(1)\, r_{21}(1) + p_{22}(1)\, r_{22}(1) = \\ &= 0{,}4 \cdot 3 + 0{,}6 \cdot (-7) = -3. \end{aligned}\right\} \tag{4.140}$$

D'une manière analogue

$$\left.\begin{aligned} h_1(2) &= p_{11}(2)\, r_{11}(2) + p_{12}(2)\, r_{12}(2) = 0{,}8 \cdot 4 + 0{,}2 \cdot 4 = 4, \\ h_2(2) &= p_{21}(2)\, r_{21}(2) + p_{22}(2)\, r_{22}(2) = \\ &= 0{,}7 \cdot 1 + 0{,}3(-19) = -5. \end{aligned}\right\} \tag{4.141}$$

Si l'on pose $v_j^0(0) = 0$ la formule (4.137) donne pour $n = 1$

$$v_1^0(1) = \max_{k(1)} \{h_1(k)\} = \max_{k(1)} \begin{cases} h_1(1) = 6, & k(1) = 1, \\ h_1(2) = 4, & k(1) = 2. \end{cases} \tag{4.142}$$

Il est clair qu'il faut opter pour $k(1) = 1$, car $v_1(1)$ acquiert alors une valeur maximale $v_1^0(1) = 6$. Si le départ a lieu à partir de l'état q_2, on a

$$v_2^0(1) = \max_{k(1)} \{h_2(k)\} = \max_{k(1)} \begin{cases} h_2(1) = -3, & k(1) = 1, \\ h_2(2) = -5, & k(1) = 2. \end{cases} \tag{4.143}$$

Là aussi il faut opter pour $k(1) = 1$; $v_2(1)$ acquiert alors la valeur maximale $v_2^0(1) = -3$.

Pour trouver les $v_i(2)$ utilisons la formule (4.137) que nous allons écrire pour $n = 2$

$$v_i^0(2) = \max_{k(2)} \left[h_i(k) + \sum_{j=1}^{2} p_{ij}(k)\, v_j^0(1) \right], \tag{4.144}$$

où $v_j^0(1)$ sont les grandeurs trouvées auparavant

$$v_1^0(1) = 6, \quad v_2^0(1) = -3. \tag{4.145}$$

Cherchons d'abord $v_1^0(2)$. Ecrivons à cette fin deux valeurs possibles de $v_1(2)$ pour différentes valeurs de k

$$\left.\begin{aligned}[v_1(2)]_{k=1} &= h_1(1) + p_{11}(1)\, v_1^0(1) + p_{12}(1)\, v_2^0(1) = \\ &= 6 + 0{,}5 \cdot 6 + 0{,}5\,(-3) = 7{,}5, \\ [v_1(2)]_{k=2} &= h_1(2) + p_{11}(2)\, v_1^0(1) + p_{12}(2)\, v_2^0(1) = \\ &= 4 + 0{,}8 \cdot 6 + 0{,}2\,(-3) = 8{,}2.\end{aligned}\right\} \quad (4.146)$$

Comme $v_1(2)$ prend sa valeur maximale pour $k = 2$, nous posons $k = 2$. Il vient $v_1^0(2) = 8{,}2$.

D'une manière analogue $v_2^0(2) = -1{,}7$, de plus la commande optimale dans ce cas $k(2) = 2$. En poursuivant ce processus, on obtient les valeurs $k_1(n)$ et $k_2(n)$, commandes optimales au départ de q_1 et de q_2 respectivement.

n	0	1	2	3	4...
$k_1(n)$	—	1	2	2	2 ...
$v_1^0(n)$	0	6	8,2	10,22	12,222 ..
$\Delta v_1^0(n) = v_1^0(n) - v_1^0(n-1)$	—	6	2,2	2,02	2,002 ...
$k_2(n)$	—	1	2	2	2 ...
$v_2^0(n)$	0	—3	—1,7	0,23	2,223 ...
$\Delta v_2^0(n) = v_2^0(n) - v_2^0(n-1)$	—	—3	—1,3	1,93	1,993 ...

On voit de la table qu'au fur et à mesure que le processus dans le système cesse d'être transitoire et devient stationnaire, l'accroissement $\Delta v_i^0(n)$ tend à une certaine valeur constante. Nous allons montrer plus loin que ce n'est pas par hasard.

Si, par exemple, jusqu'à la fin du processus il reste trois cycles, donc $n = 3$, le gain éventuel dans le cas du départ de l'état q_1 est égal à 10,22, en appliquant la commande optimale $k_1(3) = 2$.

La méthode de définition de $v_i^0(n)$ exposée ci-dessus entraîne pour de grandes valeurs de n des calculs trop volumineux. Pour définir un algorithme optimal d'un processus stationnaire avec $n \to \infty$ il est plus commode d'appliquer une autre méthode.

Avec $n \to \infty$ les probabilités d'états tendent vers les mêmes valeurs limites (finales) p_{if} indépendamment de l'état initial si le processus est totalement ergodique. Par ailleurs, l'espérance mathématique du gain pour un pas devient une grandeur constante que nous allons désigner par g. Quant au gain total $v_i(n)$ pour n pas, avec $n \to \infty$, il reçoit un accroissement constant g avec chaque

nouveau pas. Par conséquent, avec de grands n, on a pour $v_i(n)$ l'expression asymptotique suivante :

$$v_i(n) \cong ng + v_i, \tag{4.147}$$

où v_i sont les constantes qui pour les i différents peuvent être différentes. C'est précisément cette allure asymptotique qu'ont les grandeurs $v_1^0(n)$ et $v_2^0(n)$ de la table ci-dessus dans le cas de grands n.

En portant (4.147) dans (4.129), on obtient

$$ng + v_i = h_i + \sum_{j=1}^{N} p_{ij}[(n-1)g + v_j] \quad (i = 1, \ldots, N). \tag{4.148}$$

Puisque $\sum_{j=1}^{N} p_{ij} = 1$, il vient

$$g + v_i = h_i + \sum_{j=1}^{N} p_{ij} v_j \quad (i = 1, \ldots, N). \tag{4.149}$$

On dispose de N équations (4.149) pour définir $N + 1$ inconnues g, $v_1, \ldots, v_N$. Notons que l'addition d'une grandeur constante arbitraire a à tous les v_i ne change pas les équations (4.149). On en déduit que ces équations ne permettent pas de définir en général les valeurs de v_i. Elles ne peuvent donner que les différences $(v_i - v_j)$. Pourtant les valeurs mêmes de v_i ne nous intéressent pas si nous définissons comme optimal un algorithme qui en régime stationnaire maximise le gain moyen g relatif à un pas. Les équations (4.149) fournissent alors tous les renseignements nécessaires.

On peut pourtant définir la valeur de g à partir d'autres considérations. Supposons que nous avons trouvé les probabilités finales d'état p_{if} ; de plus le gain moyen par pas, comme on le voit de (4.130), est h_i. Mais puisque dans un régime stationnaire sont possibles n'importe quels q_i à probabilités p_{if}, l'espérance mathématique du gain par pas est alors définie par l'expression

$$g = \sum_{i=1}^{N} p_{if} \cdot h_i. \tag{4.150}$$

Pour appliquer cette formule il faut au préalable déterminer les probabilités finales d'états p_{if}. Nous avons montré dans ce qui précède comment il faut procéder à cet effet. Le plus simple est de définir g et v_i à partir des équations (4.149) en posant l'un des v_i, v_N par exemple, égal à zéro. On cherche alors en fait les différences $(v_i - v_N)$; pourtant ces équations donnent la valeur de g. Or, les équations (4.137) rendent possible l'obtention par itération de l'algorithme optimal $k^0(i)$ (ensemble des commandes optimales k^0 pour tous les états q_i en régime stationnaire) qui maximise le gain espéré g par pas.

En vertu de (4.137) l'algorithme optimal k^0 s'obtient en maximisant l'expression suivante par rapport à k:

$$h_i(k) + \sum_{j=1}^{N} p_{ij}(k)\, v_j(n) \tag{4.151}$$

(dans (4.137) on peut remplacer $v_i(n)$ par $v_i(n+1)$ et $v_j(n-1)$ par $v_j(n)$).

Conformément à (4.147), pour de grands n cette expression peut être remplacée par

$$h_i(k) + \sum_{j=1}^{N} p_{ij}(k)\,(ng + v_j) = h_i(k) + \sum_{j=1}^{N} p_{ij}(k)\, v_j + ng. \tag{4.152}$$

Notons que k n'intervient que dans une partie de cette expression

$$h_i(k) + \sum_{j=1}^{N} p_{ij}(k)\, v_j. \tag{4.153}$$

Pour commencer l'itération imposons-nous les valeurs initiales quelconques de v_j. On peut poser, par exemple, $v_j = 0$ pour tout j. Ayant ainsi fixé les v_j, on peut trouver pour tout i les commandes k qui maximisent l'expression (4.153). Fixons maintenant l'algorithme k obtenu et résolvons les équations (4.149) pour définir g et v_i (en posant toujours que $v_N = 0$). Après avoir calculé les nouvelles valeurs de v_i, reprenons l'expression (4.153) pour la maximiser par rapport à k et déterminer une nouvelle approximation des commandes. Après avoir trouvé k pour tous les i, fixons-les encore pour rechercher à partir de (4.149) les nouvelles valeurs de g et v_i, etc. Dans [4.7] il a été démontré que ce processus converge et donne à la limite la solution de l'équation

$$g + v_i^0 = \max_k \left[h_i(k) + \sum_{j=1}^{N} p_{ij}(k)\, v_j^0\right] \tag{4.154}$$

qui est parfaitement analogue à (4.134) pour le cas du régime stationnaire. Il ne faut pas oublier que v_i^0 et v_j^0 sont ici certains nombres qui ne dépendent pas de n.

Considérons maintenant l'exemple donné dans [4.7] pour le cas des matrices (4.138) et (4.139). Choisissons d'abord $k = 1$ pour $i = 1, 2$.

Les équations (4.149) s'écrivent dans ce cas

$$\left.\begin{aligned} g + v_1 &= 6 + 0{,}5v_1 + 0{,}5v_2, \\ g + v_2 &= -3 + 0{,}4v_1 + 0{,}6v_2. \end{aligned}\right\} \tag{4.155}$$

Posant $v_2 = 0$ trouvons

$$g = 1,\ v_1 = 10,\ v_2 = 0. \tag{4.156}$$

Maintenant il faut introduire les valeurs obtenues des v_i dans l'expression (4.153) et effectuer la maximisation par rapport à k

pour tout i. Les résultats des calculs sont donnés dans la table ci-dessous :

i	k	$h_1(k) + \sum_{j=1}^{N} p_{ij}(k) v_j$
1	1	$6+0,5\cdot 10+0,5\cdot 0=11$
	2	$4+0,8\cdot 10+0,2\cdot 0=12$
2	1	$-3+0,4\cdot 10+0,6\cdot 0=1$
	2	$-5+0,7\cdot 10+0,3\cdot 0=2$

On en déduit que pour $i = 1$ la valeur la plus avantageuse est $k = 2$; pour $i = 2$ il faut opter également pour $k = 2$.

Poursuivons le processus itératif. En fixant $k = 2$ pour tout i on obtient les équations (4.153) sous la forme

$$\left.\begin{aligned} g+v_1 &= 4+0,8v_1+0,2v_2, \\ g+v_2 &= -5+0,7v_1+0,3v_2. \end{aligned}\right\} \tag{4.157}$$

Posant $v_2 = 0$ et résolvant ces équations, il vient

$$g = 2,\ v_1 = 10,\ v_2 = 0. \tag{4.158}$$

En comparant (4.158) et (4.156) on peut voir qu'une nouvelle itération apporte un gain moyen double par rapport au précédent. L'itération peut être poursuivie. Pourtant, les valeurs des v_i fournies par (4.158) se sont avérées les mêmes que celles obtenues avec (4.150). Par conséquent, en les portant dans (4.153) et en maximisant cette expression par rapport à k, on obtient le même résultat que dans la table ci-dessus. Ainsi les valeurs des $k(i)$ obtenues à l'itération suivante coïncident avec celles de l'itération précédente. Cela signifie que les itérations ultérieures sont inutiles, que l'algorithme obtenu est optimal et qu'il vérifie la condition (4.154). Ainsi, dans l'exemple considéré, l'algorithme optimal est $k^0(1) = k^0(2) = 2$ et le gain moyen optimal par pas $g = 2$.

La méthode itérative décrite ci-dessus n'est applicable qu'aux processus stationnaires. Dans ce cas, la méthode considérée rend relativement simple la recherche de l'algorithme optimal.

Dans [4.7] est étudiée également une classe intéressante de problèmes relatifs aux systèmes discrets où la transition d'un état à l'autre est un processus aléatoire (poissonien). Pour les problèmes concernant le processus optimal dans un système continu aux entrées poissoniennes voir, par exemple, [4.21].

CHAPITRE V

SYSTÈMES OPTIMAUX À ACCUMULATION INDÉPENDANTE (PASSIVE) DE L'INFORMATION SUR L'OBJET

§ 1. Principaux problèmes de la théorie des systèmes optimaux à accumulation indépendante de l'information

Considérons toujours que l'organe de commande A reçoit une information complète sur l'opérateur F de l'objet B et sur le but final, c'est-à-dire sur la forme du critère d'optimalité Q. Quant à l'information sur la consigne $\bar{x}^*$, la perturbation $\bar{z}$, qui agit sur l'objet B, et la grandeur de sortie $\bar{x}$, elle peut être incomplète. Supposons ensuite que la quantité d'information sur les grandeurs $\bar{x}^*$, $\bar{x}$ et $\bar{z}$ peut augmenter avec le temps, ceci indépendamment du comportement de l'organe de commande A. Si ce dernier gouverne l'objet suivant une loi optimale, le système sera dit optimal, à accumulation d'information indépendante, ou passive (cf. également chapitre I).

L'accumulation de l'information peut avoir lieu dans deux cas.

a) Supposons que la grandeur $\bar{x}^*$ (ou $\bar{z}$), mesurée sans erreur, constitue un processus aléatoire plus complexe qu'un processus markovien. On peut préciser ses caractéristiques probabilistes en l'observant pendant un certain temps. Dans ce cas l'observation permet d'accumuler une information qui rend possible la prédiction du comportement futur de ce processus.

b) La grandeur $\bar{x}^*$ (ou $\bar{z}$) est mesurée avec une certaine erreur ou bien le résultat de mesure passe par un canal où le signal utile est entaché de bruits. Dans ce cas les valeurs du signal utile doivent être précisées par l'observation. Plus cette observation dure, plus l'estimation du comportement futur de $\bar{x}^*$ est précise.

Le deuxième cas est plus important.

Un exemple typique d'un système à accumulation indépendante de l'information est donné par la figure 5.1. L'information sur la valeur $\bar{x}$ de la grandeur commandée est amenée par la chaîne de réaction de la sortie de l'objet B à l'organe de commande A. A l'intérieur du circuit fermé du système les perturbations et les bruits sont absents. Pourtant, la consigne $\bar{x}^*$ est fournie à l'entrée de A par le canal H^* avec un bruit h^*. Dans ce canal au signal utile s'ajoutent

des bruits. A la sortie du canal H^* apparaît la grandeur $\bar{y}^*$, différente de $\bar{x}^*$, appliquée au lieu de $\bar{x}^*$ à l'entrée de l'organe de commande A. La tâche de l'organe de commande consiste également à séparer le signal utile du bruit ; cette tâche peut être réalisée avec assez de sûreté si l'on observe pendant une certaine période les valeurs $\bar{y}^*$. L'estimation de la valeur $\bar{x}^*$ effectuée après cette période dépend en général des valeurs $\bar{y}^*$ observées. Ainsi l'estimation de $\bar{x}^*$ et par conséquent la commande réalisée par l'organe A à l'instant courant t dépendent de la « préhistoire » de la grandeur d'entrée $\bar{y}^*(\tau)$ pour $\tau < t$. Autrement dit, la commande $\bar{u}(t)$ à l'instant t est une fonctionnelle des valeurs $y^*(\tau)$ pour $\tau < t$. Il s'ensuit qu'à la différence des organes étudiés dans les chapitres III et IV, dans le cas considéré l'organe de commande optimal A n'est plus sans inertie. Il doit être un *système dynamique* dont la grandeur de sortie à l'instant donné dépend non seulement des valeurs courantes des grandeurs d'entrée, mais encore de leurs valeurs dans le passé. L'organe de commande optimal traité dans le présent chapitre et le chapitre suivant constitue un système dynamique.

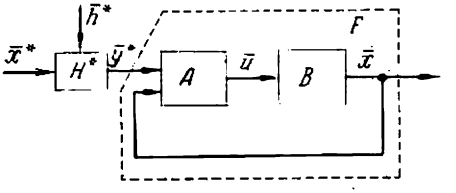


Fig. 5.1.

La figure 5.1 représente le schéma fonctionnel d'un système de commande en boucle fermée. Le problème consiste à réaliser la synthèse de l'organe de commande optimal A ou, plus précisément, à rechercher son algorithme optimal. Puisque A est maintenant un système dynamique, pour distinguer ce cas plus complexe des cas précédents bien plus simples nous allons parler de *stratégie optimale* de l'organe A.

Parfois la recherche de la stratégie optimale de l'organe A du schéma de la figure 5.1 se fait en deux étapes. A la première, on envisage le système à l'intérieur du contour en pointillé comme un seul organe F et on cherche l'algorithme optimal de cet organe. A la deuxième, on décompose F pour définir la stratégie de l'organe A, l'opérateur de l'objet B étant donné. Certes, des difficultés peuvent surgir dues à la réalisabilité ou à la grossièreté de l'organe A. Appelons *réalisable* un organe dont la grandeur de sortie peut dépendre des valeurs courantes ou antérieures des grandeurs d'entrée mais jamais des valeurs futures si seulement celles-ci ne sont pas connues d'avance. La grossièreté signifie que pour des variations suffisamment petites des paramètres de l'algorithme de l'organe de commande les variations des grandeurs de sortie ou des caractéristiques quelconques de cet organe ou du système tout entier sont aussi petites

que l'on veut. Quand on a affaire à un organe de commande qui est un système dynamique, on est obligé à imposer les conditions de réalisabilité et de grossièreté. La notion bien répandue de réalisabilité physique inclut en général les deux concepts mentionnés.

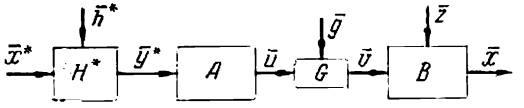


Fig. 5.2.

La figure 5.2 montre un système en boucle ouverte à accumulation indépendante de l'information. La consigne x^*, de même que sur la figure 5.1, passe par le canal H^* avec le bruit $\bar{h}^*$.

En sortant de ce canal, la consigne $\bar{y}^*$, signal utile entaché de bruit, est amenée à l'entrée de l'organe de commande A. Supposons que ce dernier agit sur l'objet B par le canal G avec le bruit $\bar{g}$. L'action réelle $\bar{v}$ qui apparaît à l'entrée de l'objet B peut donc différer de l'action $\bar{u}$ à la sortie de l'organe de commande A. Dans ce chapitre

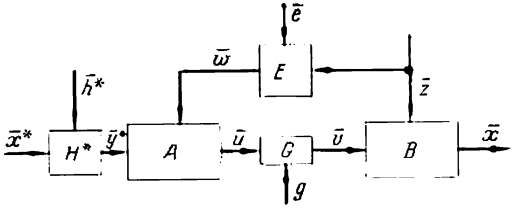


Fig. 5.3.

nous allons étudier essentiellement les systèmes en boucle ouverte dans lesquels la grandeur de sortie $\bar{x}$ n'est pas amenée à l'entrée de A.

La perturbation aléatoire $\bar{z}$ reçue par l'objet B du schéma de la figure 5.2 n'est pas mesurée. C'est pourquoi l'organe de commande A de ce schéma n'est renseigné que sur les caractéristiques probabilistes a priori de la perturbation $\bar{z}$ qui peuvent être établies, par exemple, par traitement statistique des essais réalisés dans le passé et emmagasinés par l'organe A. Dans l'expérience considérée l'organe A du schéma de la figure 5.2 ne reçoit aucun renseignement sur la perturbation $\bar{z}$.

Il se peut pourtant que la grandeur $\bar{z}$ est mesurée au cours de l'expérience et le résultat est transmis à l'entrée de l'organe A. Ce cas est montré sur la figure 5.3. Puisque les mesures s'effectuent

avec une certaine erreur, on peut représenter le résultat sous la forme de la mesure précise de la perturbation $\bar{z}$ à laquelle s'ajoute dans le canal E un bruit $\bar{e}$. La valeur $\bar{w}$ apparaissant à la sortie du canal E et amenée à l'organe de commande A diffère en général de la valeur réelle de $\bar{z}$.

Le problème de construction d'un système de commande automatique est énoncé souvent comme un problème dual. D'une part, $\bar{x}$ doit vérifier une certaine loi imposée, qui en général dépend de $\bar{x}^*$; d'autre part, $\bar{x}$ ne doit pas dépendre de la perturbation $\bar{z}$. On donne souvent à cette dernière restriction la forme d'une condition d'invariance : $\bar{x}$ doit être invariant par rapport à $\bar{z}$. Pourtant dans le cas général il est plus commode d'énoncer le critère final non pas sous la forme de deux, mais d'une seule condition, comme nous le faisons partout dans le présent ouvrage. Soit le critère final énoncé sous la forme de la condition

$$Q(\bar{x}, \bar{x}^*) = \min. \tag{5.1}$$

Dans le cas concret d'un système d'ordre n on peut appliquer un critère de la forme

$$Q = \int_0^T \sum_{i=1}^n (x_i - x_i^*)^2 \, dt. \tag{5.2}$$

Dans un système optimal le critère Q acquiert alors une valeur minimale en assurant par là une dépendance minimale entre $\bar{x}$ et la perturbation $\bar{z}$.

Si un système rend possible l'obtention d'un minimum absolu de Q, ce système peut s'appeler *idéal*. Ainsi, dans un système idéal de critère (5.2). les valeurs de x_i doivent toujours être égales à x_i^* ; dans le cas idéal Q prend donc la valeur minimale égale à zéro. Si un système idéal réalise une indépendance totale entre $\bar{x}$ et la perturbation $\bar{z}$, c'est précisément parce que $\bar{x}$ ne dépend que de $\bar{x}^*$. Bien entendu, on peut imaginer un système non idéal dans lequel $\bar{x}$ ne soit pas strictement égal à la valeur imposée $\bar{x}^*$ mais ne dépende quand même pas de $\bar{z}$.

D'après [5.1, 5.2] un système idéal donne lieu à une *invariance totale*. L'erreur $\bar{\varepsilon} = \bar{x} - \bar{x}^*$ est identiquement nulle et ne dépend pas de $\bar{x}^*$ ni de $\bar{z}$.

L'égalité de $\bar{\varepsilon}$ à zéro, c'est-à-dire l'observation de la condition d'existence d'un système idéal, est la meilleure solution possible. Si on pouvait toujours l'obtenir, la théorie des systèmes optimaux se ramènerait à la théorie de l'invariance. Pour trouver la relation entre ces deux théories différentes, il faut déterminer les classes

des systèmes pour lesquelles la condition d'invariance totale est réalisable.

La théorie de l'invariance traite, parfois sans bien les délimiter, de deux cas distincts de systèmes à mesure directe et indirecte de la perturbation $\bar{z}$. Pour le premier de ces cas le principe de compensation a été proposé en 1829 par Poncelet. Lors de la mesure directe de la perturbation $\bar{z}$ la forme de la condition de compensation est

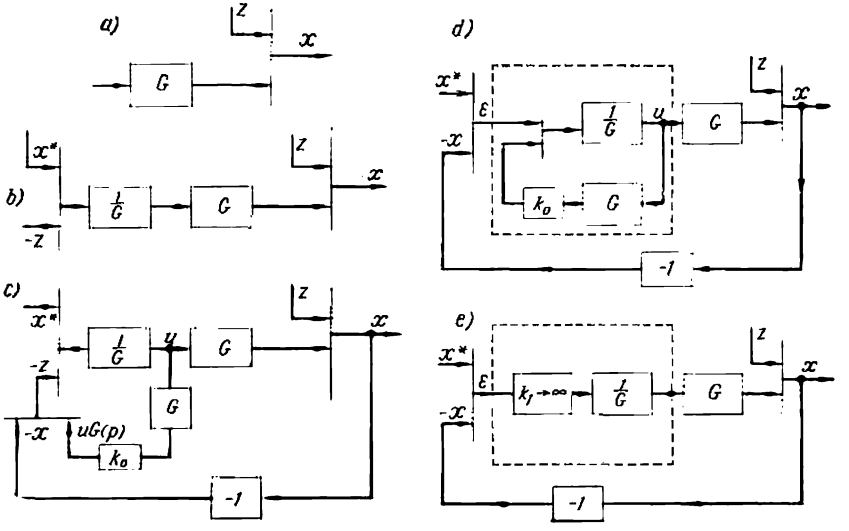


Fig. 5.4.

très simple si les mouvements de tous les circuits du système sont décrits par des équations linéaires à coefficients constants [5.2].

L'étude a porté également sur les systèmes non linéaires (cf., par exemple, [3.25], pp. 501 à 502). Dans [5.42] les méthodes variationnelles ont permis d'obtenir les conditions pour les systèmes linéaires et pour une classe très générale de systèmes non linéaires.

Dans le cas de la mesure indirecte de la perturbation $\bar{z}$ le circuit de la mesure directe est absent. Considérons à titre d'illustration un exemple élémentaire [5.3, pp. 123 à 124]. Supposons qu'à la sortie d'un système de fonction de transfert $G(p)$ apparaît la perturbation z (fig. 5.4,*a*). Ce peut être une perturbation rapportée à la sortie de l'objet, mais qui agit en d'autres points du système de commande. Considérons que le dénominateur de la fonction de transfert $G(p)$ dépend de p et que sa puissance est supérieure à celle du nominateur. C'est donc un élément avec inertie ; sa caractéristique amplitude-fréquence décroît à des fréquences suffisamment élevées. Aussi, l'objet représenté sur la figure 5.4,*a* n'est pas idéal

et lorsqu'on applique à son entrée la grandeur x^*, on obtient à sa sortie en général une autre grandeur.

Pour rendre ce système idéal (sans inertie et insensible à la perturbation z), il faut connecter en série G et un circuit de fonction de transfert $\frac{1}{G}$ (fig. 5.4,*b*). Ensuite, pour compenser la perturbation z, il faut envoyer à l'entrée du système la grandeur $(-z)$.

Un tel système peut être considéré comme un système où sont réalisées une mesure indirecte et une neutralisation de la perturbation z et éliminée l'inertie du système G par introduction du circuit $\frac{1}{G}$. En connaissant la grandeur de sortie u de l'objet ainsi que x et $G(p)$, on peut définir la grandeur z (fig. 5.4,*a*). Avec les mêmes notations u, z et x pour les transformées on a

$$z = x - uG(p). \tag{5.3}$$

En appliquant à l'entrée du système la grandeur $(-z)$, on obtient le schéma de la figure 5.4, *c*.

Le second membre de l'égalité (5.3) est élaboré (fig. 5.4,*c*) en additionnant la grandeur $(-x)$ avec la grandeur u qui passe par le bloc $G(p)$ et l'amplificateur à gain k_0 unitaire. Pour obtenir le dernier terme, il convient d'introduire l'inverseur à gain (-1).

Cependant le schéma de la figure 5.4,*c* est équivalent à celui de la figure 5.4,*d*. La fonction de transfert du circuit encadré par le pointillé est définie par l'expression

$$\frac{\frac{1}{G}}{1-k_0G\frac{1}{G}} = \frac{1}{1-k_0}\cdot\frac{1}{G} = k_1\frac{1}{G}. \tag{5.4}$$

Avec $k_0 \to 1$ le gain k_1 de ce circuit tend vers l'infini. Il en résulte le schéma représenté sur la figure 5.4,*e*.

Le bloc $\frac{1}{G}$ peut être éliminé du schéma 5.4,*e*. On obtient alors un autre schéma donné par la figure 5.5. Le gain suffisamment grand k_1 de l'amplificateur permet ici de rendre l'erreur ε aussi petite que l'on veut [5.4]. En effet, la transformée de la grandeur de sortie peut s'écrire

$$x = \frac{z}{1+k_1G} + \frac{k_1G}{1+k_1G}x^*, \tag{5.5}$$

alors que celle de l'erreur peut être mise sous la forme

$$\varepsilon = x^* - x = \frac{x^*}{1+k_1G} - \frac{z}{1+k_1G}. \tag{5.6}$$

En faisant tendre k_1 vers l'infini, on peut obtenir à la limite une erreur ε égale à zéro.

Bien entendu, ce système, en plus des oscillations forcées, donne lieu à des oscillations libres, qui peuvent s'accroître en rendant le système instable. Nous n'allons pourtant pas nous attarder aux problèmes d'instabilité. Dans [5.4] il est montré comment un schéma à réaction peut être rendu stable même lorsque le gain k est infiniment grand.

L'interprétation physique des phénomènes qui ont lieu dans le schéma de la figure 5.4,*e* ou de la figure 5.5 est assez simple. La différence $\varepsilon = x^* - x$ est fournie à l'entrée de l'amplificateur à gain k_0 infiniment grand. Dès que cette grandeur s'écarte de zéro dans

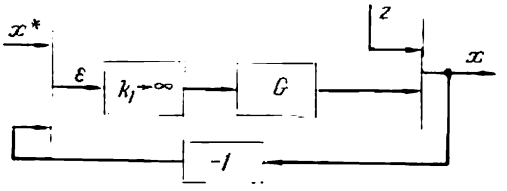


Fig. 5.5.

un sens quelconque, à la sortie de l'amplificateur apparaît une grandeur suffisamment importante qui agit sur le système en modifiant la grandeur d'entrée x de façon à réduire la différence ε à zéro.

Si dans le cas idéal on pouvait obtenir une erreur aussi petite que l'on veut ou nulle, une telle solution serait la meilleure et au lieu de la théorie des systèmes optimaux on pourrait appliquer seule la théorie de l'invariance. Malheureusement, la réalisation d'un système idéal n'est possible que si l'on pousse assez loin l'idéalisation des problèmes réels, mais alors le cadre de son application pratique devient très restreint. L'invariance peut s'obtenir dans les conditions suivantes :

a) vérification des conditions de réalisabilité et de grossièreté ;

b) limitation de l'action des perturbations aléatoires à un domaine défini (pour les détails voir plus loin) ;

c) absence de contraintes imposées directement ou indirectement à la commande et aux coordonnées du système.

Passons outre le point a) (dont l'observation est loin d'être possible dans tous les cas) et voyons comment intervient la deuxième condition. Il se trouve que même dans un schéma muni d'un amplificateur à gain important, on n'arrive pas à neutraliser *toute* perturbation z. La figure 5.6,*a* montre un tel système ; il faut pourtant tenir compte de ce que tout amplificateur réel a un coefficient de transfert $K(j\omega)$ complexe dont le module décroît brusquement aux fréquences suffisamment grandes. C'est pourquoi lorsque la perturbation z comporte des bruits d'une fréquence suffisamment élevée, sa compensation totale est impossible. Physiquement, une

cause analogue intervient également dans le schéma purement discret de la figure 5.6,*b*. Toutes les grandeurs sont examinées ici à des instants discrets $t = 0, 1, \ldots, s, \ldots$. Supposons qu'à l'instant $t = s$ la grandeur est affectée de l'indice s. Si g_s est une suite de

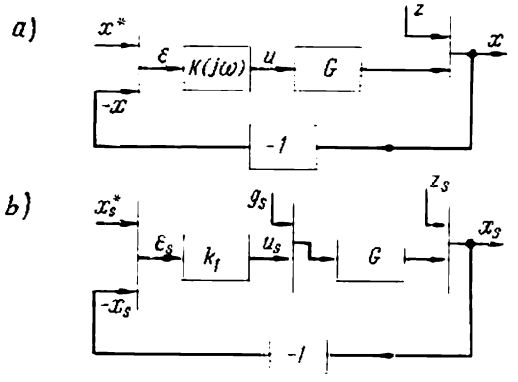


Fig. 5.6.

variables aléatoires indépendantes de loi de répartition donnée et, pour simplifier, G est posé égal à un, la mesure précise de la grandeur aléatoire z_s est impossible. En effet, la figure 5.6,*b* montre que

$$x_s = z_s + u_s + g_s, \tag{5.7}$$

d'où l'on tire

$$z_s = (x_s - u_s) - g_s. \tag{5.8}$$

Le premier terme de cette expression donne l'estimation de z_s recherchée, le deuxième terme est l'erreur de mesure. Vu que dans ce schéma z_s est mesuré indirectement et avec erreur, on ne peut obtenir une compensation précise.

Si la perturbation est appliquée à un point entre la sortie de l'amplificateur k_1 et la grandeur de sortie x, comme nous l'avons vu dans ce qui précède, dans une certaine mesure sa compensation est possible. Pourtant, dans d'autres cas tout aussi importants, lorsque la perturbation s'ajoute à la consigne x^* (fig. 5.7,*a*) ou se situe dans la chaîne de réaction en s'ajoutant à la grandeur de sortie (fig. 5.7,*b*) la situation devient bien pire. Ces cas sont parfaitement réels, car la mesure des grandeurs x^* et x se fait toujours avec une certaine erreur ; de plus les grandeurs mesurées peuvent être transmises par des chaînes d'action où elles sont entachées de bruits. Dans les schémas représentés sur la figure 5.7 l'invariance est en général impossible ; pour obtenir de meilleurs résultats, il faut donc poser et résoudre les problèmes conformément à la théorie des systèmes optimaux.

La théorie de l'invariance suppose d'ordinaire qu'il n'existe pas de contraintes directes ou indirectes sur la commande ni sur les coordonnées du système. C'est une condition très forte dont l'observation rend souvent nul l'intérêt pratique du problème. Un exemple de la contrainte indirecte peut être donné par le choix du critère d'optimalité sous la forme

$$Q = \int_0^\infty [(x - x^*)^2 + \lambda^2 u^2]\, dt = \int_0^\infty (\varepsilon^2 + \lambda^2 u^2)\, dt. \tag{5.9}$$

Si $\lambda = 0$, le système idéal s'obtient avec $\varepsilon = 0$ et $Q = 0$. Toutefois, en ajoutant le terme $\lambda^2 u^2$ à l'expression sous le signe somme, on restreint l'effet intégral de l'action u en interdisant ses valeurs trop grandes. Dans ce cas le minimum de Q n'est plus nul (cf. chapitre III) et pour un système optimal la grandeur ε n'est pas identiquement nulle. Le même résultat s'obtient si l'on applique le critère intégral généralisé (cf. chapitre I).

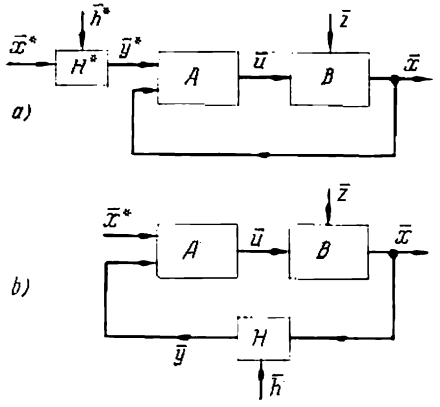


Fig. 5.7.

Le rôle le plus important revient pourtant aux contraintes directes dont le sens est traduit par le fait que la commande et les coordonnées du système *ne peuvent* ou *ne doivent pas* dépasser certaines limites admissibles. Par conséquent, la grandeur de sortie de l'amplificateur des figures 5.4, 5.5, 5.6 ne doit pas dépasser la valeur admissible. On en tire immédiatement que l'obtention de la valeur $\varepsilon \equiv 0$ est impossible. Or, dans la grande majorité des problèmes l'idéalisation est raisonnable tant que les grandeurs peuvent être limitées par des valeurs admissibles.

Il convient de noter que les procédés fondés sur la compensation lors de la mesure directe (cf. fig. 5.3) sont un auxiliaire précieux dans l'étude des systèmes. Ces procédés, qui ont à leur base le principe de Poncelet, ne constituent pas l'essentiel de la théorie de l'invariance mais forment une branche à part étudiée depuis de nombreuses années.

Il est possible de réaliser une invariance partielle, par exemple, assurer l'absence d'erreur dans un processus stable. Les procédés correspondants (introduction des intégrateurs ou autres procédés) sont également bien connus et n'ont aucun rapport direct avec la

théorie de l'invariance. De plus, avec des perturbations aléatoires le calcul d'un régime stable se ramène aux problèmes typiques de la théorie des systèmes optimaux lorsque l'erreur est différente de zéro et la tâche consiste à minimiser une des mesures de cette erreur (la variance, par exemple).

La diminution de l'erreur par introduction des amplificateurs à gain important et à réaction négative profonde, comme dans le schéma de la figure 5.5, permet d'obtenir une petite erreur permanente tout en assurant la stabilité du système [5.4]. Les schémas de ce type ont trouvé une très large application, par exemple, dans la simulation électronique [3.25], [5.5]. Ce domaine constitue également une branche particulière.

Pour ce qui est de la branche principale de la théorie de l'invariance, étudiée pour la première fois dans [5.6], étant formellement correcte, elle traite des systèmes superidéalisés. Dans cette théorie la mesure indirecte de la perturbation et l'élimination de l'inertie s'obtiennent par compensation et application d'une réaction positive dans le but d'obtenir un gain infini. L'influence des bruits aléatoires et des contraintes n'est pas prise ici en considération. Les systèmes de ce type ne vérifient pas, en outre, la condition de grossièreté [5.6]. Souvent le rôle des contraintes revient à des nonlinéarités de types variés; la théorie de l'invariance ne peut donc étudier non plus les systèmes de ce type. Par conséquent, cette théorie ne peut être appliquée que dans un cadre très étroit qui ne promet pas de s'étendre; au contraire, il diminue toujours du fait qu'en technique on observe un progrès continu dans le sens d'une précision plus poussée et d'une meilleure utilisation des ressources des systèmes d'où le fonctionnement du système aux régimes limites, la prise en considération des contraintes et des perturbations aléatoires.

Certains ouvrages traitent de ce qu'on appelle l'« invariance à ε près ». Cela signifie que, l'invariance totale étant impossible, ces ouvrages se consacrent à l'étude des problèmes relatifs à une erreur minimale non nulle et pouvant être quelconque, même aussi grande que l'on veut. Cette formulation du problème ne se distingue déjà plus de celle donnée par la théorie des systèmes optimaux. Dans ce domaine présente un grand intérêt une classe particulière des problèmes où il faut déterminer les conditions de faible sensibilité du système aux variations des perturbations. L'étude de la sensibilité peut s'effectuer lors de l'analyse des systèmes. Mais si cette notion est introduite dans un problème de synthèse, elle peut figurer sous la forme d'un critère d'optimalité ou de contraintes et donc faire partie des problèmes de la théorie des systèmes optimaux.

Les systèmes en boucle ouverte examinés dans ce chapitre présentent un grand intérêt pour la technique des communications. C'est précisément dans ce domaine que la théorie des systèmes opti-

maux à accumulation de l'information a trouvé sa première application. Or, cette théorie présente également un grand intérêt pour les systèmes de commande automatique (cf. chapitres V et VI).

Malgré la grande diversité des problèmes, ces dernières années on a vu se préciser dans la théorie des communications trois groupes principaux suivants de théories.

1) *Théories de corrélation.* Les problèmes de ce groupe ont ce trait particulier qu'on ne connaît pas toutes les caractéristiques probabilistes mais seulement les fonctions de corrélation des processus aléatoires. Il se trouve que si l'on se borne aux systèmes linéaires et au critère d'optimalité sous la forme du minimum de l'erreur quadratique moyenne, ces données suffisent pour construire, par exemple, un filtre optimal. Dans le présent ouvrage nous n'examinons pas ce groupe de théories. La littérature relative à ce domaine est donnée dans la bibliographie du chapitre I.

Les théories de corrélation peuvent être appliquées également au calcul des systèmes non linéaires mais dans ce cas leurs possibilités sont limitées.

2) *Théorie d'information.* Cette théorie a pour objet les problèmes de transmission des communications d'un point de vue très général. Sa tâche principale consiste à rechercher un code optimal, les propriétés du canal de transmission et la nature statistique de la source des communications étant données. Bien que ce problème ait plusieurs solutions intéressantes, à la différence des questions de transmission et de stockage de l'information, les méthodes de la théorie relatives au traitement de l'information ne sont pas encore au point.

3) *Théorie des décisions statistiques.* Il semble que cette théorie soit la plus générale et applicable à n'importe quels processus des systèmes de communications et de commande tant en boucle ouverte que fermée. Dans notre exposé nous nous basons sur les concepts de cette théorie.

L'une des sources de la théorie des décisions statistiques est la statistique mathématique, branche importante du calcul des probabilités. Vers la fin du XIXe siècle et la première moitié du XXe siècle, dans ce domaine on a vu apparaître la théorie d'estimations des paramètres (R. Fischer, H. Cramer et, plus tard, le savant soviétique Y. Linnik [5.7 à 5.9]) et la théorie de vérification d'hypothèses (J. Neumann, E. Pearson, A. Kolmogorov).

La théorie des jeux a débuté indépendamment de la statistique mathématique. Cette discipline dont les fondements ont été posés par le mathématicien français E. Borel, a été développée par J. von Neumann et autres (1928) tout en pénétrant au cours de son évolution sous des formes différentes dans les théories des opérations, de programmation linéaire (dont les principes ont été établis par le mathématicien soviétique L. Kantorovitch en 1938) et de pro-

grammation dynamique. L'objet principal de la théorie des jeux est le jeu à deux ou à plusieurs personnes, les règles du jeu et notamment les gains éventuels des participants étant connus. Le problème consiste à élaborer pour chaque participant une stratégie optimale, c'est-à-dire la méthode de recherche de la « meilleure » solution. Même dans la pire des situations une telle stratégie doit fournir le coup le plus avantageux.

Notre ouvrage n'exposant pas la théorie des jeux, le lecteur qui s'y intéresse peut s'adresser aux ouvrages [5.10 à 5.14].

La réunion et le développement de plusieurs idées énoncées ci-dessus ont conduit en 1948-1950 le mathématicien américain A. Wald à construire une discipline mathématique statistique générale qui a reçu le nom de théorie des fonctions de décision [5.15]. Il semble que plusieurs idées de R. Bellman qui s'est consacré à la programmation dynamique sont dans une certaine mesure liées aux principes de la théorie des fonctions de décision.

Indépendamment de toutes ces disciplines mathématiques, la théorie des communications a donné lieu à des problèmes très généraux de caractère statistique. Le premier ouvrage consacré à la construction et aux propriétés d'un récepteur optimal a été publié en 1946 par le savant soviétique V. Kotelnikov [5.16]. Un récepteur optimal rend minimale la probabilité d'erreur. V. Kotelnikov a donné à sa théorie le nom de stabilité potentielle aux perturbations. Par la suite des recherches consacrées aux méthodes optimales de réception des ondes radio ont été poursuivies intensément en U.R.S.S. et à l'étranger.

Vers 1950 les branches mathématique et technique de la théorie des décisions statistiques ont fusionné. Les travaux du savant américain D. Middleton et autres ont montré que les méthodes de la théorie des fonctions de décision peuvent être appliquées avec succès à la résolution des problèmes relatifs au meilleur procédé de réception des signaux en présence des bruits [1.12, 1.13, 5.17-5.22]. Nous montrons dans ce chapitre et les chapitres suivants que les méthodes de la théorie des décisions statistiques permettent de résoudre également les problèmes de la théorie de la commande automati que.

Considérons d'abord les problèmes principaux de la théorie des décisions statistiques appliqués aux systèmes de communication. Sur le schéma fonctionnel d'un système de communication (fig. 5.8) E désigne l'émetteur, CC la chaîne de communication, R le récepteur. Le signal transmis par l'émetteur est noté $s(t)$, celui reçu par le récepteur $x(t)$. Dans la chaîne de communication le bruit $n(t)$ vient entacher le signal $s(t)$. Par conséquent

$$x(t) = x[s(t), n(t)]. \tag{5.10}$$

Dans un cas particulier, par exemple,

$$x(t) = s(t) + n(t). \tag{5.11}$$

Supposons que certains paramètres du signal transmis sont inconnus (s'ils étaient tous connus, il n'y aurait aucun besoin de transmettre ce signal). Mais il se peut qu'on connaît leur répartition probabiliste a priori ; cette répartition peut s'obtenir, par exemple, par traitement des données statistiques accumulées pendant une longue période révolue. Supposons encore qu'on connaît le mode de combinaison du signal et du bruit dans la chaîne de communication, c'est-à-dire le type de la formule (5.10). Admettons de plus que les caractéristiques probabilistes du bruit $n(t)$ sont connues.

Voici la formulation du problème : soit le signal $x(t)$ appliqué à l'entrée du récepteur R en un intervalle de temps fini de $t = 0$ à

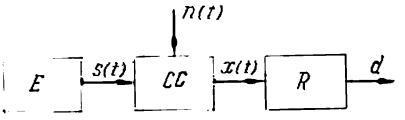


Fig. 5.8.

$t = T$, c'est-à-dire $0 \leqslant t \leqslant T$. Sur la base de la réalisation reçue du signal $x(t)$ il faut déterminer, par un procédé dans un certain sens optimal, les paramètres inconnus du signal transmis $s(t)$. Supposons que cette solution s'obtient automatiquement dans le récepteur R et est fournie à sa sortie sous la forme d'un signal d (fig. 5.8). Trouver l'algorithme du récepteur optimal.

Pour concrétiser le problème, il faut indiquer le critère d'optimalité et les paramètres du signal $s(t)$ qui doivent être déterminés. Suivant la formulation du problème, divers types de problèmes et de théories sont possibles.

1) *Théorie des décisions bialternatives.* Soit un seul paramètre A inconnu. Le signal $s(t)$ peut s'écrire, par exemple,

$$s(t) = A \sin(\omega_0 t + \varphi_0), \tag{5.12}$$

où ω_0 et φ_0 sont donnés. Dans la théorie des décisions bialternatives on considère le cas où le paramètre inconnu ne peut prendre que deux valeurs possibles A_1 et A_2 avec des probabilités a priori $p(A_1)$ et $p(A_2)$ respectivement. Dans un cas particulier $A_1 \neq 0$, $A_2 = 0$. Ce problème s'appelle problème de la détection du signal.

2) *Théorie des décisions multialternatives.* Dans cette théorie on pose que le paramètre inconnu A peut prendre r valeurs possibles différentes $A_1, A_2, \ldots, A_r$ avec des probabilités a priori $p(A_1)$, $p(A_2), \ldots, p(A_r)$ respectivement.

3) *Théorie d'estimation des paramètres.* Soit le paramètre inconnu A qui peut prendre dans un certain domaine $\Omega(A)$ une valeur quelconque avec une densité de probabilité a priori $P(A)$. La théorie de

l'estimation des paramètres permet de construire un récepteur optimal qui donne à la sortie la meilleure dans un certain sens estimation de la valeur du paramètre A du signal transmis $s(t)$ sur la base de la réalisation reçue $x(t)$.

4) *Théorie d'estimation des processus.* Nous avons posé ci-dessus qu'au cours de la transmission le paramètre A du signal transmis est constant. Or, il se peut que A soit une fonction du temps : $A = A(t)$. Le problème consiste à définir de façon optimale la forme et les paramètres de cette fonction.

Les problèmes énoncés ci-dessus ne sont pas les plus généraux. Dans le cas général le signal transmis peut avoir plusieurs paramètres inconnus que nous allons noter $a_1, a_2, \ldots, a_m$:

$$s(t) = s(a_1, a_2, \ldots, a_m, t). \tag{5.13}$$

Chacun de ces paramètres peut prendre plusieurs valeurs possibles ou même, dans un certain domaine, un nombre infini de valeurs. Si l'on introduit le vecteur des paramètres A aux coordonnées $a_1, a_2, \ldots, a_m$

$$A = (a_1, a_2, \ldots, a_m), \tag{5.14}$$

les problèmes d'estimation de plusieurs paramètres peuvent être énoncés sous la même forme que les problèmes relatifs à un seul paramètre, il faut considérer seulement que A n'est pas un scalaire, mais un vecteur.

Voici quelques traits caractéristiques des formulations des problèmes précédents.

1) On reçoit d'abord toute la réalisation x dans l'intervalle $0 \leqslant t \leqslant T$, puis on prend une décision d sur le paramètre A. Certes, une autre formulation est également possible : on commence à estimer les valeurs de A dès l'instant $t = 0$ en précisant l'estimation progressivement sur la base de l'information fournie au récepteur.

2) Dans la majorité des cas nous allons considérer que les probabilités a priori des valeurs possibles du paramètre A sont connues. Les problèmes de ce type sont appelés *problèmes de Bayes.*

Pour résoudre les problèmes examinés dans ce chapitre, on peut faire appel à des notions géométriques : l'espace du bruit, l'espace du signal et l'espace des observations. Introduisons d'abord la notion d'espace du bruit. Supposons que les mesures sont prises seulement à des instants discrets $t_1, t_2, \ldots, t_k$. Nous ne nous intéressons donc qu'aux valeurs du bruit $n_1, n_2, \ldots, n_k$ à ces instants précis (fig. 5.9). Introduisons le vecteur du bruit N aux coordonnées cartésiennes $n_1, n_2, \ldots, n_k$:

$$N = (n_1, n_2, \ldots, n_k). \tag{5.15}$$

Ce rayon vecteur dans un espace de dimension k, appelé espace du bruit, est représenté sur la figure 5.10. Dans notre exemple cet

espace est tridimensionnel. Désignons par $P(N)$ la densité de probabilité du vecteur N, c'est-à-dire la densité de probabilité conjointe des coordonnées $n_1, n_2, \ldots, n_k$:

$$P(N) = P(n_1, n_2, \ldots, n_k). \tag{5.16}$$

Définissons la probabilité avec laquelle l'extrémité du vecteur N rejoint le volume infiniment petit $d\Omega(N)$ de l'espace du bruit, par

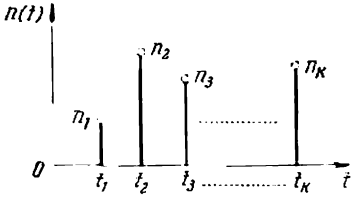


Fig. 5.9.

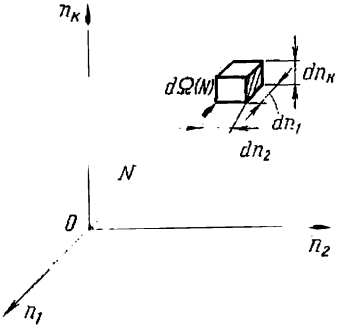


Fig. 5.10.

exemple, un parallélépipède infiniment petit de volume

$$d\Omega(N) = dn_1\, dn_2 \ldots, dn_k \tag{5.17}$$

représenté sur la figure 5.10. Il est clair que cette probabilité est égale à *)

$$P(N)\, d\Omega(N) = P(n_1, n_2, \ldots, n_k)\, dn_1\, dn_2 \ldots dn_k. \tag{5.18}$$

La probabilité totale qu'a l'extrémité du vecteur N d'atteindre un certain point de l'espace du bruit $\Omega(N)$ étant égale à l'unité, l'intégrale de tous les volumes infiniment petits $d\Omega(N)$ de cet espace est égale à l'unité:

$$\int\limits_{\Omega(N)} P(N)\, d\Omega(N) = 1, \tag{5.19}$$

l'intégration s'étendant à tout le domaine $\Omega(N)$ des valeurs possibles de N, c'est-à-dire à tout l'espace du bruit.

L'espace des paramètres est associé au vecteur des paramètres A défini par (5.14) de la même façon que l'espace du bruit est lié au vecteur N. Cet espace est un espace de dimension m aux coordonnées cartésiennes $a_1, a_2, \ldots, a_m$. Si

$$P(A) = P(a_1, a_2, \ldots, a_m)$$

*) Cf. la figure 2.6 et les raisonnements qui s'y rapportent.

est la densité de probabilité du vecteur A, le produit $P(A)\,d\Omega(A)$ est la probabilité avec laquelle l'extrémité du vecteur A atteint le volume infiniment petit $d\Omega(A)$ de l'espace du signal.

Introduisons enfin l'*espace des observations* ou, autrement dit, l'espace du signal reçu. En considérant ses valeurs discrètes x_1, $x_2, \ldots, x_k$ on peut (cf. également (2.97)) introduire le vecteur

$$X = (x_1, x_2, \ldots, x_k). \tag{5.20}$$

L'espace à k dimensions de ce vecteur est précisément l'espace des observations. Si l'on connait la densité de probabilité $P(X) = P(x_1, x_2, \ldots, x_k)$, la probabilité avec laquelle l'extrémité du vecteur X rejoint le volume infiniment petit $d\Omega(X)$ de l'espace des observations $\Omega(X)$ est définie par le produit $P(X)\,d\Omega(X)$. Il est évident également que

$$\int\limits_{\Omega(X)} P(X)\,d\Omega(X) = 1. \tag{5.21}$$

Il convient de noter que $P(N)$ et $P(A)$ figurent parmi les données du problème, alors que $P(X)$ ne l'est pas mais peut être défini au cours de la résolution du problème.

Ainsi supposons que l'on connaisse :

a) le type de la fonction qui caractérise le signal transmis

$$s(t) = s(t, A);$$

b) la densité de probabilité a priori $P(A)$ du vecteur des paramètres A, ou les probabilités $p(A_i)$ des valeurs particulières A_i si les A_i forment un ensemble fini ou dénombrable ;

c) la densité de probabilité $P(N)$ du vecteur du bruit N ;

d) le mode de combinaison du signal s et du bruit n dans le canal de transmission :

$$x = x[s, n].$$

Supposons qu'on reçoit les valeurs discrètes $x_1, x_2, \ldots, x_n$ du signal x et qu'on définit la réalisation concrète du vecteur X.

Au fond, toute l'information sur A fournie par l'expérience est concentrée dans la densité de probabilité a posteriori de A. C'est une densité de probabilité conditionnelle $P(A \mid X)$ ou, autrement dit, la densité de probabilité de A sous la condition que le vecteur observé X soit fixé. Sous une forme développée

$$P(A \mid X) = P(a_1, a_2, \ldots, a_m \mid x_1, x_2, \ldots, x_k). \tag{5.22}$$

Pour définir la densité de probabilité a posteriori on peut se servir de la formule de Bayes (2.13). Considérons d'abord le cas où le vecteur A peut prendre un nombre fini r de valeurs possibles $A_1, A_2, \ldots, A_r$ avec les probabilités a priori $p(A_1), p(A_2), \ldots, p(A_r)$ respectivement. Soit $P(X \mid A)\,d\Omega(X)$ la probabilité conditionnelle du fait que l'extrémité du vecteur X se trouve dans

le volume $d\Omega\ (X)$ de l'espace des observations à la condition que le vecteur A soit donné. Trouvons la probabilité de l'événement complexe qui consiste dans ce que le vecteur des paramètres ait la valeur A et que l'extrémité du vecteur X rejoigne le volume $d\Omega\ (X)$. D'après la formule du produit des probabilités, cette probabilité peut être exprimée de deux façons. Elle vaut

$$p\ (A)\ [P\ (X \mid A)\ d\Omega\ (X)] = p\ (A \mid X)\ [P\ (X) \cdot d\Omega\ (X)], \qquad (5.23)$$

où $P\ (X)$ est la densité de probabilité inconditionnelle du vecteur X. On en tire

$$p(A \mid X) = \frac{p(A) P(X \mid A)}{P(X)}. \qquad (5.24)$$

La fonction $P\ (X \mid A)$ s'appelle fonction de vraisemblance. Connaissant la probabilité a priori $p\ (A_i)$ et ayant obtenu le signal transmis X, on peut déterminer à l'aide de (5.24) la probabilité a posteriori

$$p\ (A_i \mid X) = \frac{p\ (A_i)\ P\ (X \mid A_i)}{P\ (X)}. \qquad (5.25)$$

Le dénominateur de cette expression peut être défini de la manière suivante : en additionnant les expressions

$$p\ (A_j)\ P\ (X \mid A_j) = P\ (X)\ p\ (A_j \mid X) \qquad (5.26)$$

pour $j = 1, \ldots, r$, on obtient

$$\sum_{j=1}^{r} p\ (A_j)\ P\ (X \mid A_j) = P\ (X) \sum_{j=1}^{r} p\ (A_j \mid X), \qquad (5.27)$$

d'où, compte tenu du fait que

$$\sum_{j=1}^{r} p\ (A_j \mid X) \quad 1, \qquad (5.28)$$

on a

$$P\ (X) = \sum_{j=1}^{r} p\ (A_j)\ P\ (X \mid A_j). \qquad (5.29)$$

En portant cette expression dans (5.25) on tombe en définitive sur

$$p\ (A_i \mid X) = \frac{p\ (A_i)\ P\ (X \mid A_i)}{\sum_{j=1}^{r} p\ (A_j)\ P\ (X \mid A_j)} \quad (i = 1, \ldots, r). \qquad (5.30)$$

Une formule analogue se déduit également pour le cas de la répartition continue du vecteur A de densité a priori $P\ (A)$. Soit $P\ (A \mid X)$ la densité de probabilité de A pour X donné, c'est-à-dire une densité de probabilité a posteriori. La probabilité inconditionnelle de ce que l'extrémité du vecteur A se trouve dans le volume $d\ \Omega\ (A)$ de l'espace des paramètres est égale à $P\ (A)\ d\Omega\ (A)$, alors qu'avec X

fixé, la probabilité conditionnelle est égale à $P(A \mid X)\, d\Omega(A)$. La probabilité d'un événement complexe pour que les extrémités des vecteurs A et X se trouvent dans les volumes $d\Omega(A)$ et $d\Omega(X)$ des espaces respectifs est égale à $P(A, X)\, d\Omega(A)\, d\Omega(X)$, où $P(A, X)$ est la densité de probabilité conjointe correspondante. D'après le théorème du produit des probabilités on a

$$P(A, X)\, d\Omega(A)\, d\Omega(X) = [P(A)\, d\Omega(A)]\, [P(X \mid A)\, d\Omega(X)] = \\ = [P(X)\, d\Omega(X)]\, [P(A \mid X)\, d\Omega(A)]. \quad (5.31)$$

On en tire

$$P(A)\, P(X \mid A) = P(X)\, P(A \mid X) \quad (5.32)$$

et la formule de Bayes pour la densité de probabilité a posteriori

$$P(A \mid X) = \frac{P(A)\, P(X \mid A)}{P(X)}. \quad (5.33)$$

Pour obtenir le dénominateur de cette formule, multiplions les deux membres de (5.32) par $d\Omega(A)$ et intégrons sur tout le domaine $\Omega(A)$ des valeurs possibles du vecteur A. Il vient :

$$\int_{\Omega(A)} P(A)\, P(X \mid A)\, d\Omega(A) = P(X) \int_{\Omega(A)} P(A \mid X)\, d\Omega(A). \quad (5.34)$$

L'intégrale du second membre de cette expression est égale à l'unité ; c'est pourquoi

$$P(X) = \int_{\Omega(A)} P(A)\, P(X \mid A)\, d\Omega(A). \quad (5.35)$$

Cette expression doit être portée dans (5.33). On obtient alors la formule de Bayes sous sa forme définitive :

$$P(A \mid X) = \frac{P(A)\, P(X \mid A)}{\int_{\Omega(A)} P(A)\, P(X \mid A)\, d\Omega(A)} \quad (5.36)$$

Ces formules sont celles de la *méthode de probabilité a posteriori* qui consiste à choisir A de façon à maximiser la probabilité a posteriori $p(A \mid X)$. Utilisant cette méthode, on obtient comme décision d à la sortie du récepteur optimal la valeur *la plus probable* de A.

Si la répartition du paramètre A est continue, d'après la méthode considérée il faut choisir la valeur de A telle qu'elle maximise la densité de probabilité a posteriori $P(A \mid X)$.

Il faut pourtant noter que l'application de cette méthode n'est possible que dans le cas des probabilités a priori connues de diverses valeurs de A. S'il n'en est pas ainsi, on peut appliquer une autre méthode décrite dans ce qui suit.

Considérons la fonction de vraisemblance $P(X \mid A)$. Si le vecteur X est donné, cette fonction ne dépend que de A :

$$P(X \mid A) = L(A). \tag{5.37}$$

La méthode la plus simple d'estimation de A est associée à la fonction de vraisemblance et s'appelle *méthode du maximum de vraisemblance*. Cette méthode, proposée par le mathématicien anglais R. Fischer, peut être énoncée sous la forme de la règle suivante:

La valeur la plus vraisemblable du paramètre A est celle qui maximise la fonction de vraisemblance $L(A)$.

Cette règle, adoptée comme postulat, n'est justifiée que par des raisonnements logiques.

Pour évaluer le paramètre A d'après la méthode du maximum de vraisemblance on procède comme suit: supposons que l'observation fournit un signal sous la forme du vecteur X. Dans le cas d'une répartition discrète de A, écrivons toutes les valeurs possibles de la fonction de vraisemblance

$$L(A_1),\ L(A_2),\ \ldots,\ L(A_r).$$

Après avoir comparé ces valeurs, choisissons A_j telle que

$$L(A_j) \geqslant L(A_k) \quad (k = 1, \ldots, r). \tag{5.39}$$

La valeur A_j choisie d'après cette méthode est dite *la plus vraisemblable.*

Cette méthode se généralise naturellement aux problèmes à répartition continue du paramètre A. On choisit une valeur de A telle qu'elle maximise la valeur de $P(X \mid A) = L(A)$.

§ 2. Eléments de la théorie générale des décisions statistiques

Soit le paramètre a du signal transmis $s = af(t)$ qui peut prendre $m + 1$ valeurs possibles $a_0, a_1, a_2, \ldots, a_m$ avec des probabilités a priori $p_0, p_1, p_2, \ldots, p_m$ respectivement.

Supposons que le récepteur reçoit un signal entaché du bruit $n(t)$. On demande de déterminer d'après la courbe $x(t)$, observée dans l'intervalle $0 \leqslant t \leqslant T$, la valeur du paramètre a dans le signal transmis. Quelle que soit la méthode de résolution, elle consiste à diviser l'espace des observations en $m + 1$ domaines $\Gamma_0, \Gamma_1, \ldots, \Gamma_m$ qui correspondent aux solutions $a = a_0$, $a = a_1$, etc. Dans le cas considéré on peut appliquer la méthode du maximum de vraisemblance, du maximum de probabilité a posteriori, la méthode de Kotelnikov (minimum de probabilité inconditionnelle de l'erreur) ou une variante de la méthode de Neumann-Pearson. Dans ce dernier cas on s'impose les probabilités de certaines erreurs, alors que la somme des probabilités des autres erreurs est minimisée.

La généralisation ultérieure est donnée par les problèmes d'estimation des paramètres. Soit le paramètre a du signal $s(t) = af(t)$ qui a une répartition a priori continue de densité $P(a)$. Il faut rechercher d'après l'échantillon reçu $X = (x_1, x_2, \ldots, x_k)$ des valeurs du signal $x(t)$ ou d'après la réalisation de la fonction $x(t)$ dans l'intervalle $0 \leqslant t \leqslant T$ la valeur du paramètre a dans le signal transmis. Soit, par exemple, $x_i = s_i + n_i$ $(i = 1, \ldots, k)$, alors que les valeurs du bruit n forment une suite de variables aléatoires indépendantes d'espérance mathématique nulle et de variance σ^2. La fonction de vraisemblance s'écrit alors pour une valeur quelconque du paramètre a

$$L(a) = P(X \mid a) = \frac{1}{(2\pi)^{k/2}\sigma^k} \exp\left\{-\frac{\sum_{i=1}^{k}[x_i - af(t_i)]^2}{2\sigma^2}\right\}. \quad (5.40)$$

En se guidant sur la méthode du maximum de vraisemblance, choisissons une valeur de a telle que la fonction de vraisemblance $L(a)$ soit maximale, ce qui est équivalent à la condition

$$F(a) = \sum_{i=1}^{k}[x_i - af(t_i)]^2 =$$

$$= \sum_{i=1}^{k} x_i^2 - 2a\sum_{i=1}^{k} x_i f(t_i) + a^2\sum_{i=1}^{k}[f(t_i)]^2 = \min. \quad (5.41)$$

Egalons à zéro la dérivée $\frac{dF(a)}{da}$:

$$-2\sum_{i=1}^{k} x_i f(t_i) + 2a\sum_{i=1}^{k}[f(t_i)]^2 = 0. \quad (5.42)$$

L'observation de la condition (5.42) minimise la fonction $F(a)$. On en tire l'estimation $a = a^*$ du maximum de vraisemblance:

$$a^* = \frac{\sum_{i=1}^{k} x_i f(t_i)}{\sum_{i=1}^{k}[f(t_i)]^2}. \quad (5.43)$$

Pour passer au cas continu, on peut multiplier le numérateur et le dénominateur de (5.43) par Δt. Avec $\Delta t \to 0$ et $k \to \infty$ les sommes se transforment en intégrales et la formule (5.43) devient

$$a^* = \frac{\int_0^T x(t) f(t)\, dt}{\int_0^T [f(t)]^2\, dt} = \frac{1}{e_0}\int_0^T x(t) f(t)\, dt, \quad (5.44)$$

où

$$e_0 = \int_0^T [f(t)]^2 \, dt. \tag{5.45}$$

La grandeur e_0 peut s'appeler énergie spécifique du signal. L'expression (5.44) montre que la définition de a^* se ramène à la détection synchrone, c'est-à-dire à la formation du produit $x(t) f(t)$, puis à l'intégration cohérente qui consiste à obtenir l'intégrale de ce produit.

Dans le cas général le problème d'estimation des processus consiste à définir la forme du signal transmis $s(t)$. On peut la définir approximativement si l'on donne, par exemple, une suite des échantillons $s_1, s_2, \ldots, s_k$, où k est un nombre suffisamment grand. Il s'agit d'estimer ces grandeurs.

Introduisons le vecteur

$$S = (s_1, s_2, \ldots, s_k) \tag{5.46}$$

aux coordonnées cartésiennes $s_1, s_2, \ldots, s_k$ dans un espace de dimension k. Appelons cet espace *espace du signal*. Introduisons une densité de probabilité a priori $P(S)$ de ce vecteur. Alors $P(S)\, d\Omega(S)$ est la probabilité pour que l'extrémité du vecteur S rejoigne le volume infiniment petit $d\Omega(S)$ de cet espace.

Dans le cas particulier d'estimation des paramètres, le signal $s(t)$ s'écrit $s(t, A)$, où $A = (a_1, \ldots, a_r)$ est le vecteur des paramètres ; de plus, le type de la relation entre s d'une part et t et A de l'autre est connu.

Le problème de réception peut être considéré comme un problème de décision, alors que le récepteur lui-même est au fond un organe de décision. Sa tâche est de fournir à la sortie la *décision* sur la forme du signal transmis. Si on doit estimer un processus discret, c'est-à-dire le vecteur $S = (s_1, s_2, \ldots, s_k)$, il faut donner les estimations $d_1, \ldots, d_k$ pour chacune des grandeurs $s_1, \ldots, s_k$. Par conséquent, dans le cas général, la solution est donnée par le vecteur

$$D = (d_1, d_2, \ldots, d_k). \tag{5.47}$$

Introduisons l'*espace des décisions* qui est un espace de dimension k du vecteur D, construit de la même façon que celui du signal S. Ses coordonnées ne sont plus les valeurs du signal s_i mais leurs estimations d_i effectuées par le récepteur. Si le problème consiste à estimer le vecteur des paramètres A, la décision recherchée porte sur les coordonnées $a_1, \ldots, a_r$. En d'autres termes, la sortie du récepteur doit fournir r décisions $d_1, \ldots, d_r$ qui correspondent au vecteur D de dimension r. Dans ce cas l'espace des décisions est identique à l'espace des paramètres A.

A chaque réalisation du signal reçu, c'est-à-dire à tout point X de l'espace des observations, il faut faire correspondre une décision D, c'est-à-dire un vecteur D dans l'espace des décisions. Toute règle de décision est une règle de transformation des points de X-espace (espace des observations) en points de D-espace (espace des décisions). Dans la théorie des décisions statistiques cette règle qui constitue l'algorithme de l'organe de réception s'appelle *stratégie* du récepteur.

Deux types de stratégies sont possibles. Dans le premier, à chaque X fixé correspond un point défini de D-espace. Une telle stratégie s'appelle *déterministe* ou *pure*. Dans le deuxième type de stratégies, à chaque X fixé correspond une certaine densité de probabilité $\Delta(D \mid X)$ des points de D-espace. Cela signifie que le choix de la décision s'effectue d'une manière aléatoire; pourtant, la loi statistique qui régit ce choix dépend de la valeur observée X. Les stratégies du deuxième type s'appellent *aléatoires*.

Le deuxième type de stratégies est plus général que le premier. En effet, une stratégie pure peut être considérée comme un cas limite d'une stratégie aléatoire lorsque la variance des points de D-espace associés au vecteur observé X_0 tend vers zéro. Par ailleurs, les points de D-espace se concentrent pratiquement dans le voisinage immédiat d'un certain point $D_0(X_0)$ qui correspond à X_0 d'après une loi uniforme définie. A la limite, la densité de probabilité $\Delta(D \mid X_0)$ peut être décrite par l'expression

$$\Delta(D \mid X_0) = \delta[D - D_0(X_0)]. \tag{5.48}$$

Ici δ désigne l'impulsion unité égale à zéro aux points $D \neq D_0(X_0)$ et infinie en $D_0(X_0)$. De plus, l'intégrale

$$\int_D \delta[D - D_0(X_0)]\, d\Omega(D) \tag{5.49}$$

sur tout D-espace est égale à l'unité du fait que dans la formule (5.48) les premier et second membres sont des densités de probabilité.

Ainsi, donner la fonction $\Delta(D \mid X)$ est au fond équivalent à donner une stratégie aléatoire. Dans le cas particulier où Δ se transforme en δ-fonction, la stratégie devient pure.

La fonction $\Delta(D \mid X)$ porte le nom de *fonction de décision* (cf. [5.15]). C'est précisément cette fonction qui définit la stratégie de l'organe de décision qu'il faut trouver. Dans la théorie des décisions statistiques le problème consiste à déterminer la fonction de décision optimale, le critère d'optimalité étant associé aux erreurs de la décision.

Si le signal transmis est S et la décision adoptée D, il se peut que D ne soit pas une estimation correcte de S. Supposons qu'une décision incorrecte entraîne une certaine « pénalité » dont la valeur

peut être appréciée par la *fonction de pertes*, ou *de pénalisation*, W dépendant en général de S et de D :

$$W = W(S, D). \tag{5.50}$$

Dans le cas d'une décision correcte les pertes ne pouvant pas être supérieures à celles dues à une décision incorrecte, il vient

$$W(S, D = S) \leqslant W(S, D \neq S). \tag{5.51}$$

Pour une décision correcte, la pénalité est choisie d'ordinaire la plus petite, c'est-à-dire l'inégalité (5.51) est observée strictement.

Diverses fonctions vérifient cette condition. Par exemple, lors de l'estimation des coordonnées $S_1, \ldots, S_k$ du vecteur S on peut choisir

$$W(S, D) = \text{const}\,[(s_1 - d_1)^2 + \ldots + (s_k - d_k)]^2 \tag{5.52}$$

ou

$$W(S, D) = \text{const}\,[\,|(s_1 - d_1|\, + \ldots + |\, s_k - d_k\,|], \tag{5.53}$$

etc. Parfois la fonction du type

$$W(S, D) = \begin{cases} -\infty, & S = D, \\ 1, & S \neq D \end{cases} \tag{5.54}$$

est appelée simple, ou élémentaire. Cette fonction peut être exprimée par la formule

$$W(S, D) = 1 - \delta(S - D), \tag{5.55}$$

où δ désigne la fonction δ.

Supposons que le signal à transmettre S soit fixé. Sous cette condition la décision D est en général aléatoire. En effet, premièrement, le vecteur $\bar{X}$ dont la valeur guide le choix de la décision est aléatoire par suite de la présence du bruit N. Deuxièmement, X étant fixé, D peut être aléatoire si le récepteur réalise une stratégie aléatoire. Mais si D est aléatoire, la fonction de pertes $W(S, D)$ est également aléatoire et ne peut servir de mesure de qualité du fonctionnement d'un organe de décision. Il est logique de prendre pour telle mesure l'espérance mathématique de la fonction de pertes qui porte le nom de *risque*.

Appelons risque conditionnel $r(S, \Delta)$ l'espérance mathématique de la fonction de pertes $W(S, D)$ pour un signal S fixé et une certaine fonction de décision Δ fixée. D'après la définition de l'espérance mathématique, il faut prendre la moyenne de $W(S, D)$ sur différentes valeurs de D. Désignons par $P(D \mid S)$ la densité de probabilité conditionnelle de D pour S donné. Il vient

$$r(S, \Delta) = M\{W \mid S\} = \int\limits_{\Omega(D)} W(S, D)\, P(D \mid S)\, d\Omega. \tag{5.56}$$

$\Omega(D)$ est ici le domaine des valeurs possibles de D et $d\Omega$ son volume infiniment petit. Nous désignerons désormais par $d\Omega$ le volume infiniment petit de n'importe quel domaine sur lequel on réalise l'intégration.

Connaissant les conditions de l'expérience, c'est-à-dire les caractéristiques probabilistes du bruit et le mode de combinaison du signal et du bruit dans le canal de transmission, on peut obtenir la densité de probabilité $P(X \mid S)$. Si l'on connaît la loi du fonctionnement de l'organe de décision, c'est-à-dire la fonction $\Delta(D \mid X)$, la fonction $P(D \mid S)$ qui figure dans la formule (5.56) est définie par l'expression

$$P(D \mid S) = \int\limits_{\Omega(X)} \Delta(D \mid X)\, P(X \mid S)\, d\Omega. \tag{5.57}$$

Dans le cas général les signaux transmis sont inconnus. On ne sait que leur densité de probabilité a priori $P(S)$. Il faut donc prendre la moyenne du risque conditionnel r sur tout le domaine $\Omega(S)$ des signaux possibles. Le résultat d'une telle opération noté R s'appelle risque *total*, ou *moyen*. La grandeur R est l'espérance mathématique du risque conditionnel r pour des observations distinctes lorsque la source envoie des signaux S avec une densité de probabilité $P(S)$. L'expression de R s'écrit

$$R = M\{r\} = \int\limits_{\Omega(S)} r(S, \Delta)\, P(S)\, d\Omega. \tag{5.58}$$

Dans le cas considéré $d\Omega$ désigne un volume infiniment petit de l'espace des signaux $\Omega(S)$.

Maintenant on peut énoncer le problème de recherche de la stratégie optimale de l'organe de décision comme un problème de définition d'une fonction de décision $\Delta(D \mid X)$ telle qu'elle minimise le risque moyen R.

Tous les problèmes de la théorie des décisions statistiques sont des cas particuliers du problème général de minimisation du risque. Cette thèse peut être illustrée par un exemple tiré de la théorie des décisions bialternatives à paramètre unique a. Dans ce cas, dans les formules respectives S est remplacé par a, et D par la grandeur scalaire d. Supposons que le paramètre a peut prendre deux valeurs possibles a_1 et a_0 aux probabilités a priori p_1 et p_0 respectivement; d, également, ne peut alors prendre que deux valeurs possibles

$$d = \begin{cases} d_1 & (\text{décision } a = a_1), \\ d_0 & (\text{décision } a = a_0). \end{cases} \tag{5.59}$$

Soit dans l'exemple considéré la fonction de pertes

$$W = \begin{cases} 0, & d = a, \\ 1, & d \neq a. \end{cases} \tag{5.60}$$

Ecrivons l'expression du risque conditionnel $r(a, \Delta)$. Supposons que les probabilités conditionnelles des décisions d_1 et d_0, pour a fixé, soient $p(d_1 \mid a)$ et $p(d_0 \mid a)$. Dans l'exemple considéré l'intégrale (5.56) est remplacée par la somme

$$r(a, \Delta) = W(a, d_0)\, p(d_0 \mid a) + W(a, d_1)\, p(d_1 \mid a). \tag{5.61}$$

Notamment,

$$r(a_1, \Delta) = W(a_1, d_0)\, p(d_0 \mid a_1) + W(a_1, d_1)\, p(d_1 \mid a_1) = \\ = W(a_1, d_0)\, p(d_0 \mid a_1) = p(d_0 \mid a_1) \tag{5.62}$$

et

$$r(a_0, \Delta) = W(a_0, d_0)\, p(d_0 \mid a_0) + W(a_0, d_1)\, p(d_1 \mid a_0) = \\ = W(a_0, d_1)\, p(d_1 \mid a_0) = p(d_1 \mid a_0). \tag{5.63}$$

Ensuite, l'intégrale de la formule (5.58) du risque moyen R est remplacée dans l'exemple considéré par la somme

$$R = M\{r\} = r(a_1, \Delta)\, p_1 + r(a_0, \Delta)\, p_0, \tag{5.64}$$

du fait que p_1 et p_0 sont les probabilités a priori des valeurs a_1 et a_0 respectivement. En portant (5.62) et (5.63) dans (5.64), on a

$$R = p_0 \cdot p(d_1 \mid a_0) + p_1 \cdot p(d_0 \mid a_1). \tag{5.65}$$

Mais $p(d_1 \mid a_0)$ est la probabilité d'une « fausse alarme » :

$$p(d_1 \mid a_0) = \alpha. \tag{5.66}$$

D'une manière parfaitement analogue on obtient $p(d_0 \mid a_1)$ qui est la probabilité de la décision $a = a_0$ alors qu'en vérité $a = a_1$, c'est-à-dire la probabilité d'une « fausse fin d'alerte » :

$$p(d_0 \mid a_1) = \beta. \tag{5.67}$$

Après avoir porté (5.66) et (5.67) dans (5.65) cette dernière devient

$$R = p_0\alpha + p_1\beta = q, \tag{5.68}$$

ce qui signifie que dans le cas considéré le risque moyen R coïncide avec la probabilité inconditionnelle q de l'erreur. C'est pourquoi, la condition $R = \min$ pour une stratégie optimale traduit dans le cas considéré la minimisation de la probabilité inconditionnelle q de l'erreur. On aboutit ainsi au critère de Kotelnikov.

On peut montrer que choisissant autrement la fonction de pertes W on arrive à des conditions qui coïncident avec d'autres cri-

tères d'optimalité et, notamment, avec le critère de Neumann-Pearson.

Le problème de minimisation du risque total R énoncé ci-dessus ne peut être posé que si l'on connaît la densité de probabilité a priori $P(S)$ du signal S. Le problème en question porte le nom de Bayes et la fonction $\Delta(D \mid X)$ fournie par sa résolution s'appelle *décision de Bayes.*

Toutefois, dans de nombreux problèmes de grand intérêt pratique la densité de probabilité $P(S)$ est inconnue. Dans ce cas il faut appliquer la formule du risque conditionnel r (5.56) dont le calcul n'impose pas la connaissance de la fonction $P(S)$. La théorie des décisions statistiques propose la méthode de raisonnement « pessimiste » suivante : on trouve d'abord le « pire » des signaux S^* tel qu'il maximise le risque conditionnel r (pour Δ donné). Cette condition s'écrit sous la forme analytique suivante :

$$r(S^*, \Delta) = \max_{(S)} r(S, \Delta). \tag{5.69}$$

Comparons maintenant les diverses fonctions de décision et sélectionnons celle d'entre elles (appelons-la Δ^*) qui rende minimal le « pire » risque conditionnel $r(S^*, \Delta)$. Ainsi :

$$r(S^*, \Delta^*) = \min_{(\Delta)} r(S^*, \Delta) = \min_{(\Delta)} \max_{(S)} r(S, \Delta). \tag{5.70}$$

La décision Δ^* est dite décision du *minimax* et la stratégie qui lui correspond stratégie *optimale du minimax.*

C'est une stratégie qui dans le pire des cas donne un résultat meilleur (c'est-à-dire une valeur plus faible du risque r) que toute autre stratégie dans le pire des cas pour cette dernière.

On aurait tort pourtant du point de vue pratique de se guider par ces considérations si les pires conditions du fonctionnement sont très rares. L'approche du minimax n'est donc pas du tout la seule possible. Depuis Laplace on recommande parfois, dans le cas où la répartition a priori $P(S)$ est inconnue, de la considérer comme uniforme, c'est-à-dire d'adopter a priori que dans le domaine $\Omega(S)$ toutes les valeurs de S sont équiprobables. Fait curieux à noter qu'une telle approche conduit souvent à la stratégie du minimax.

Il convient de noter également que la stratégie optimale du minimax Δ n'est au fond nullement une stratégie uniformément optimale pour tout signal S. Pour différents types de signaux, elle peut donner un résultat pire que les autres stratégies.

Les notions de risque, de stratégies de Bayes et du minimax exposées dans ce qui précède se présentent par leur autre côté si l'on introduit encore un espace, l'*espace du risque.* Considérons cet espace pour le cas le plus simple de la théorie des décisions bialternatives lorsqu'il se transforme en un plan. Etudions d'abord les

stratégies pures D_i. Supposons qu'on a adopté une stratégie quelconque D_1. Soit $r(a_0, D_1)$ le risque conditionnel lié à l'application de cette stratégie dans le cas où $a = a_0$. Rappelons que cette expression est l'espérance mathématique de la fonction de pertes à la condition que $a = a_0$ et que l'on applique la stratégie D_1. Mais si $a = a_1$, l'application de la stratégie D_1 est liée au risque conditionnel $r(a_1, D)$.

Portons les valeurs du risque conditionnel $r(a_0, D)$ et $r(a_1, D)$ sous la forme de coordonnées cartésiennes dans un plan que nous appellerons plan du risque (fig. 5.11). A la stratégie D_1 correspond le point D_1 de coordonnées $r(a_0, D_1)$ et $r(a_1, D_1)$ dans le plan du risque. Pour une autre stratégie D_2, les risques conditionnels $r(a_0, D_2)$ et $r(a_1, D_2)$ auront en général d'autres valeurs. Dans le plan du risque cette stratégie correspond à un autre point D_2.

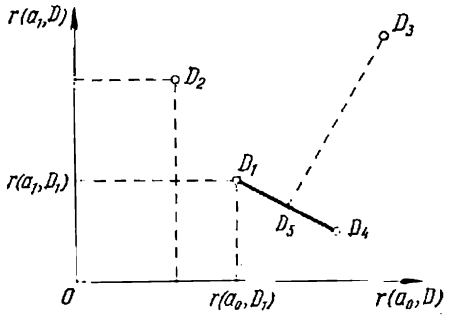


Fig. 5.11.

Plus le risque est petit, meilleure est la stratégie. Toutefois, une comparaison directe des stratégies D_1 et D_2 dans le plan du risque ne rend pas toujours possible le choix de l'une d'elles. Par exemple, la valeur du risque $r(a_1, D_1)$ pour la stratégie D_1 est plus faible que la valeur respective $r(a_1, D_2)$ pour la stratégie D_2 (cf. fig. 5.11). Mais par contre le risque $r(a_0, D_2)$ associé à la stratégie D_2 est plus faible. Il est clair que la stratégie D_3 doit être rejetée du fait que pour cette stratégie les deux risques conditionnels $r(a_0, D_3)$ et $r(a_1, D_3)$ sont plus grands que pour la stratégie D_1.

Jusqu'à présent nous n'avons figuré dans un plan du risque que les stratégies pures. Or, les stratégies aléatoires peuvent également être représentées dans le plan du risque. En effet, imaginons que nous choisissons soit la stratégie D_1 avec la probabilité q_1 soit la stratégie D_4 avec la probabilité q_2. Désignons cette stratégie aléatoire par D_5'. Le risque conditionnel r quand $a = a_0$ est égal alors à la valeur moyenne

$$r(a_0, D_5') = q_1 r(a_0, D_1) + q_2 r(a_0, D_4). \tag{5.71}$$

D'une manière analogue

$$r(a_1, D_5') = q_1 r(a_1, D_1) + q_2 r(a_1, D_4). \tag{5.72}$$

On voit aisément que le point D_5' de coordonnées $r(a_0, D_5')$ et $r(a_1, D_5')$ repose sur la droite reliant D_1 et D_4.

Comme $q_1 + q_2 = 1$ et q_1 varie de 0 à 1, le point D_6' peut prendre une position quelconque sur le segment D_1D_4 en se déplaçant de D_4 à D_1. Les stratégies pour lesquelles on emploie d'une façon aléatoire les différentes stratégies pures sont dites parfois *mixtes*. On en tire qu'à tout point du segment entre deux points matérialisant des stratégies pures correspond une stratégie mixte.

Réunissons maintenant les points D_5' et D_3 par une droite (trait interrompu de la figure 5.11). A tout point du segment en pointillé correspond alors une certaine stratégie mixte. Elle s'obtient en appliquant les stratégies D_3 et D_5' avec des probabilités définies

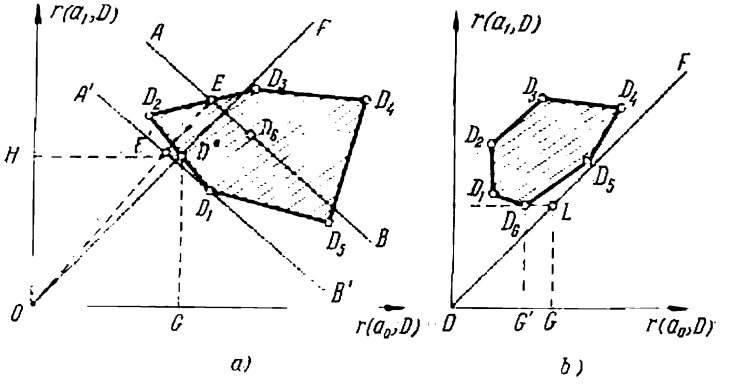


Fig. 5.12.

(ce qui conduit finalement à l'application, avec certaines probabilités, de trois stratégies pures D_1, D_4 et D_3). Réunissant de la même façon des points quelconques des stratégies obtenues auparavant, on peut voir que n'importe quel point intérieur du polygone convexe formé par les stratégies pures primaires, par exemple par D_1, D_2, D_3, D_4, D_5 de la figure 5.12, correspond à une stratégie mixte quelconque. Sur la figure 5.12 ce polygone est hachuré. Si un point associé à une stratégie pure primaire quelconque D_6 se trouve à l'intérieur du polygone, il ne participe pas à sa construction.

Quelle est l'interprétation de la stratégie de Bayes dans le plan du risque? Se donnant les probabilités a priori p_0 et p_1 des valeurs a_0 et a_1 respectivement, on obtient le risque moyen R d'après la formule

$$R = p_1 r(a_1, D) + p_0 r(a_0, D). \tag{5.73}$$

La droite $R = \text{const}$ dans le plan du risque est une droite de pente $\frac{p_0}{p_1}$. Si p_0 et p_1 sont donnés, la pente de la droite $R = \text{const}$ est également donnée. Admettons que c'est la droite AB sur la

figure 5.12, *a*. On voit sans peine que la valeur du risque R pour les points de cette droite, c'est-à-dire le premier membre de l'équation (5.73), est proportionnelle à la longueur de la perpendiculaire OE abaissée de l'origine des coordonnées sur la droite AB. C'est le cas de la stratégie D_6, par exemple. Cependant, la figure 5.12, *a* montre qu'on peut obtenir une stratégie meilleure pour laquelle le risque R sera plus petit. A cette fin il faut déplacer la droite AB parallèlement à elle-même dans le sens de l'origine des coordonnées. A la position extrême $A'B'$ de cette droite correspond la plus petite longueur possible OE' de la perpendiculaire et, par conséquent, la plus petite valeur possible du risque total R. La droite AB ne peut plus s'approcher de l'origine des coordonnées car toutes les stratégies possibles D se situent dans les limites du polygone hachuré. Une seule stratégie D_1 correspond à la position extrême $A'B'$ de la droite; c'est précisément la stratégie optimale. La construction rend évident qu'une stratégie optimale correspond au sommet du polygone. Et puisque tous les sommets correspondent à des stratégies pures, on en tire une conclusion de grand intérêt.

Une stratégie de Bayes optimale est une stratégie pure.

Cette conclusion est vraie également pour des problèmes bien plus généraux que le problème examiné.

Toutefois une stratégie optimale n'est pas toujours une stratégie pure. Lorsque le problème est considéré en se guidant par le principe du minimax, il faut chercher une stratégie D^* capable d'assurer une valeur minimale du maximum du risque conditionnel $r(a, D^*)$, c'est-à-dire permettant de choisir la moindre des deux valeurs maximales possibles de $r(a_0, D^*)$ et de $r(a_1, D^*)$. Pour déterminer la stratégie optimale du minimax menons la bissectrice OF de l'angle compris entre les axes de coordonnées. Supposons que la bissectrice OF coupe le polygone des stratégies. Soit un certain point qui se déplace de l'origine suivant la droite OF. L'endroit de sa première rencontre avec le contour du polygone correspond précisément à la stratégie D^*. En effet, pour ce point les deux valeurs du risque $OG = r(a_0, D^*)$ et $OH = r(a_1, D^*)$ sont égales entre elles et inférieures au risque maximal r pour tout autre point du polygone. Ainsi, pour le point D_1 le risque maximal $r(a_1, D_1)$ est plus grand que OG.

Dans le cas général, la bissectrice OF ne coupe pas le polygone à l'un de ses sommets. C'est pourquoi la stratégie D^* est mixte et non pas pure. On en tire la conclusion importante:

Une stratégie optimale du minimax est une stratégie aléatoire.

Dans un cas particulier, la stratégie optimale du minimax peut être une stratégie pure. Supposons, par exemple, que la bissectrice OF ne coupe pas le polygone $D_1D_2D_3D_4D_5$ (fig. 5.12, *b*). Associons au point L, qui se déplace de l'origine des coordonnées suivant la bissectrice, la droite verticale LG et la droite horizontale LH. La

stratégie optimale correspond au point du polygone qui rencontre le premier l'une de ces droites : sur la figure 5.12, *b* c'est D_6.

Toutes ces conclusions s'étendent à une large classe de problèmes. Par exemple, la généralisation à la théorie des décisions multialternatives consiste à remplacer le plan du risque par un espace de dimension $m+1$ aux coordonnées $r(a_0, D)$, $r(a_1, D)$, . . ., $r(a_m, D)$. Les méthodes et les résultats, décrits ci-dessus pour un plan, s'étendent complètement au cas d'un espace multidimensionnel.

§ 3. Application de la théorie des décisions statistiques aux systèmes de commande automatique

L'application de la théorie des décisions statistiques aux systèmes de commande en boucle ouverte ou aux systèmes réductibles aux systèmes de ce type ne diffère pas en principe de l'application aux problèmes de la théorie des communications. Pourtant, la littérature consacrée à ces problèmes est moins riche. Citons [5.23]

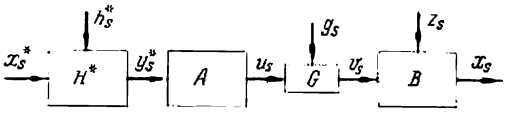


Fig. 5.13.

qui examine la théorie du système optimal d'identification des images basée sur la théorie des décisions statistiques ; [5.24] qui, avant d'exposer la théorie des systèmes en boucle fermée, étudie un système en boucle ouverte ; dans [5.25] sont exposés certains concepts fondamentaux de la théorie des décisions statistiques et donné un exemple très simple d'application de cette théorie au calcul d'un système de commande, ce dernier se ramenant en fait à un système en boucle ouverte.

Considérons le système de commande en boucle ouverte représenté sur la figure 5.13 [5.24]. Pour simplifier, supposons que dans ce schéma les grandeurs sont des scalaires et des fonctions du temps discret $t = 0, 1, \ldots, s, \ldots, n$, où n est fixé. Soit la consigne x^* qui s'écrit

$$x_s^* = x_s^*(s, \overline{\lambda}), \tag{5.74}$$

où $\overline{\lambda}$ est un vecteur aux coordonnées aléatoires :

$$\overline{\lambda} = (\lambda_1, \ldots, \lambda_q). \tag{5.75}$$

Ainsi, x_s^* est un processus aléatoire discret. Voici quelques exemples concrets d'un tel processus :

$$x_s^* = \lambda_1, \tag{5.76}$$

$$x_s^* = \lambda_1 \exp\{-\lambda_2 s\}, \tag{5.77}$$

$$x_s^* = \sum_{i=1}^{q} \lambda_i f_i(s), \tag{5.78}$$

où $f_i(s)$ sont des fonctions connues.

Dans le cas particulier de (5.76) le processus aléatoire dégénère en variable aléatoire λ_1.

Soit la densité de probabilité a priori $P(\bar{\lambda})$ du vecteur $\bar{\lambda}$.

Supposons que la densité de probabilité $P(h_s^*)$ du bruit h^* dans une chaîne sans inertie H^*, ainsi que le mode de combinaison du signal x^* et du bruit h^* dans cette chaîne soient connus. On peut calculer alors la probabilité conditionnelle $P(y_s^* \mid x_s^*)$. Cette fonction est la même pour tout s, car la densité $P(h_s^*)$ pour le bruit est considérée comme invariable. Admettons en outre que x_s^*, h_s^* et toutes les autres actions extérieures sont indépendantes. Soit, enfin h_i^* $(i = 1, \ldots, n)$ une suite de variables indépendantes.

Dans le cas le plus simple l'objet commandé B est sans mémoire et son opérateur est donné par l'expression

$$x_s = F(z_s, v_s), \tag{5.79}$$

où F est la fonction connue, tandis que la perturbation z_s dépend du temps discret s et du vecteur aléatoire $\bar{\mu}$:

$$z_s = z_s(s, \bar{\mu}). \tag{5.80}$$

Le vecteur $\bar{\mu}$ compte m coordonnées :

$$\bar{\mu} = (\mu_1, \ldots, \mu_m). \tag{5.81}$$

Soit $P(\bar{\mu})$ la densité de probabilité a priori. La grandeur d'entrée v_s du système B se forme à la sortie de la chaîne d'action sans inertie G. En passant par cette chaîne la commande u_s est entachée du bruit g_s dont la densité de probabilité $P(g_s)$ est invariable.

Les valeurs g_s constituent une suite de variables aléatoires indépendantes.

Désignons par $W_s(s, x_s^*, x_s)$ la fonction de pertes élémentaire qui correspond à l'instant discret s. Supposons que la fonction de pertes totale W est égale à la somme des fonctions de pertes élémentaires

$$W = \sum_{s=0}^{n} W_s(s, x_s^*, x_s). \tag{5.82}$$

Le problème consiste à définir la stratégie optimale de l'organe de commande A. Dans le cas général, on considère que cette stratégie est aléatoire. La densité de probabilité de la commande u_s

à l'instant $t = s$ est désignée par $\Gamma_s(u_s)$. En général, cette grandeur doit dépendre de toute l'information d'entrée emmagasinée par l'organe de commande A en un temps $t = 0, 1, \ldots, s-1$, c'est-à-dire des valeurs $y_0^*, y_1^*, \ldots, y_{s-1}^*$.

Introduisons les vecteurs temporels

$$\left.\begin{aligned} \vec{y}_s^* &= (y_0^*, y_1^*, \ldots, y_s^*), \\ \vec{x}_s^* &= (x_0^*, x_1^*, \ldots, x_s^*) \quad (s = 0, 1, \ldots, n). \end{aligned}\right\} \tag{5.83}$$

Les coordonnées du vecteur $\vec{y}_s^*$ constituent une suite des valeurs du scalaire y^* aux instants $t = 0, 1, \ldots, s$. Des suites analogues des valeurs de la grandeur x^* forment un vecteur temporel $\vec{x}_s^*$. Nous désignerons les vecteurs temporels par une flèche au-dessus de la lettre correspondante.

Comme la stratégie optimale, c'est-à-dire la densité de probabilité optimale $\Gamma_s(u_s)$, dépend du vecteur $\vec{y}_{s-1}^*$, on peut la représenter sous la forme d'une densité de probabilité conditionnelle $\Gamma_s(u_s \mid \vec{y}_{s-1}^*)$. Le problème consiste à rechercher les distributions optimales Γ_s pour tous les instants $t = 0, 1, \ldots, s, \ldots, n$. Il faut sélectionner les fonctions Γ_s $(s = 0, 1, \ldots, n)$ de façon à minimiser le risque moyen R, c'est-à-dire l'espérance mathématique de la fonction de pertes W:

$$R = M\{W\} = M\{\sum_{s=0}^{n} W_s\} = \sum_{s=0}^{n} M\{W_s\} = \sum_{s=0}^{n} R_s, \tag{5.84}$$

où la grandeur

$$R_s = M\{W_s(s, x_s^*, x_s)\} \tag{5.85}$$

peut se nommer *risque élémentaire*. Le risque élémentaire R_s correspond à l'instant discret $t = s$.

Désignons par $P(x_s \mid u_s)$ la densité de probabilité conditionnelle de x_s avec u_s fixé. Soit, ensuite, $P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*)$ la densité de probabilité conditionnelle du vecteur $\vec{y}_{s-1}^*$ pour $\vec{x}_{s-1}^*$ fixé. Les h_i^* étant indépendants, on a

$$P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*) = P(y_0^*, y_1^*, \ldots, y_{s-1}^* \mid \vec{x}_{s-1}^*) = \\ = \prod_{i=0}^{s-1} P(y_i^* \mid \vec{x}_{s-1}^*) = \prod_{i=0}^{s-1} P(y_i^* \mid x_i^*). \tag{5.86}$$

Puisque le vecteur $\vec{x}_{s-1}^*$ dépend de s et de $\bar{\lambda}$, l'expression (5.86) dépend également de s et de $\bar{\lambda}$. La figure 5.14 représente le schéma dans lequel l'action des facteurs aléatoires de chaque bloc du système

est prise en considération par introduction de densités de probabilité conditionnelles.

Désignons par $\Omega\,(x_s,\ v_s,\ u_s,\ \vec{y}^*_{s-1})$ le domaine de variation des paramètres $x_s,\ v_s,\ u_s,\ y^*_0,\ \ldots,\ y^*_{s-1}$. On peut l'imaginer sous la forme d'un espace de dimension $(s+3)$ des coordonnées cartésiennes constituées par tous les paramètres indiqués ci-dessus. Soit $d\Omega\,(x_s,\ v_s,\ u_s,\ \vec{y}^*_{s-1}) = dx_s dv_s du_s dy^*_0 \ldots dy^*_{s-1}$. Dans ce qui suit, le volume infiniment petit d'un domaine quelconque est noté par $d\Omega$; on comprend toujours de quel domaine il s'agit.

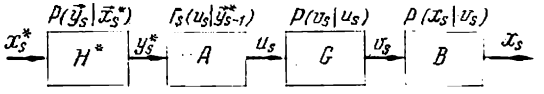


Fig. 5.14.

Ecrivons d'abord l'expression du risque *conditionnel élémentaire* r_s en entendant par là la valeur du risque élémentaire pour le vecteur fixe $\vec{x}_s^*$ ou, ce qui revient au même, pour le vecteur fixe $\overline{\lambda}$. Il vient

$$r_s = M\,\{W_s \mid \vec{x}_s^*\} = \int\limits_{\Omega(x_s)} W_s\,(s,\ x_s^*,\ x_s)\,P\,(x_s \mid \vec{x}_s^*)\,d\Omega. \qquad (5.87)$$

La moyenne de W_s est prise ici par rapport à la variable aléatoire x_s. Puisqu'on considère que $\vec{x}_s^*$ est fixé, en calculant la moyenne il faut utiliser la probabilité conditionnelle de x_s pour x_s^* fixé, c'est-à-dire la fonction $P\,(x_s \mid \vec{x}_s^*)$.

Etant donné que x_s dépend de la grandeur d'entrée v_s de l'objet B, on peut, connaissant la caractéristique probabiliste a priori de la perturbation z_s, trouver $P\,(x_s \mid v_s)$ (fig. 5.14). Mais la grandeur v_s est elle-même une variable aléatoire et la loi de sa répartition dépend en définitive de $\vec{x}_s^*$, c'est-à-dire est la fonction $P\,(v_s \mid \vec{x}_s^*)$. La relation entre ces densités s'écrit

$$P\,(x_s \mid \vec{x}_s^*) = \int\limits_{\Omega(v_s)} P\,(x_s \mid v_s)\,P\,(v_s \mid \vec{x}_s^*)\,d\Omega. \qquad (5.88)$$

En portant (5.88) dans (5.87), on a :

$$r_s = \int\limits_{\Omega(x_s,\ v_s)} W_s\,(s,\ x_s^*,\ x_s)\,P\,(x_s \mid v_s)\,P\,(v_s \mid \vec{x}_s^*)\,d\Omega. \qquad (5.89)$$

Développons l'expression (5.89).

Connaissant $P\,(v_s \mid u_s)$ on peut trouver la densité de probabilité

$$P\,(v_s \mid \vec{x}_s^*) = \int\limits_{\Omega(u_s)} P\,(v_s \mid u_s)\,P\,(u_s \mid \vec{x}_s^*)\,d\Omega. \qquad (5.90)$$

D'une façon parfaitement analogue on obtient l'expression

$$P(u_s \mid \vec{x}_s^*) = \int\limits_{\Omega(\vec{y}_{s-1}^*)} \Gamma_s(u_s \mid \vec{y}_{s-1}^*) P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*)\, d\Omega, \qquad (5.91)$$

où l'intégration se fait sur le domaine $\Omega(y_0^*, \ldots, y_{s-1}^*) = \Omega(\vec{y}_{s-1}^*)$.

En portant (5.91) dans (5.90) et (5.90) dans (5.89) on aboutit à l'expression pour le risque conditionnel élémentaire

$$r_s = M\{W_s \mid \vec{x}_s^*\} = \int\limits_{\Omega(x_s,\, v_s,\, u_s,\, \vec{y}_{s-1}^*)} W_s(s, x_s^*, x_s)\, P(x_s \mid v_s) \times$$

$$\times P(v_s \mid u_s)\, \Gamma_s(u_s \mid \vec{y}_{s-1}^*)\, P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*)\, d\Omega. \qquad (5.92)$$

Le facteur $P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*)$ sous le signe somme dépend du vecteur $\bar{\lambda}$. Désignons par $\Omega(\bar{\lambda})$ le domaine de variation du vecteur $\bar{\lambda}$. Dans les différentes expériences ce vecteur peut prendre des valeurs différentes. C'est pourquoi le risque conditionnel élémentaire r_s a également, suivant l'expérience, des valeurs différentes et constitue de ce fait une variable aléatoire. Le risque élémentaire moyen R_s peut s'obtenir en calculant la moyenne du risque conditionnel r_s par rapport à $\bar{\lambda}$:

$$R_s = \int\limits_{\Omega(\bar{\lambda})} r_s P(\bar{\lambda})\, d\Omega. \qquad (5.93)$$

En portant (5.92) dans (5.93), il vient :

$$R_s = \int\limits_{\Omega(x_s,\, u_s,\, v_s,\, \vec{y}_{s-1}^*,\, \bar{\lambda})} W_s(s, x_s^*, x_s)\, P(x_s \mid v_s)\, P(v_s \mid u_s) \times$$

$$\times \Gamma_s(u_s \mid \vec{y}_{s-1}^*)\, P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*)\, P(\bar{\lambda})\, d\Omega. \qquad (5.94)$$

Le problème consiste à choisir les fonctions Γ_s telles qu'elles minimisent les valeurs de R_s et par conséquent le risque total $P = \sum\limits_{s=0}^{n} R_s$.

Considérant le problème énoncé, on voit sans peine qu'il ne se distingue en rien au fond des problèmes de la théorie des communications examinés au § 2 de ce chapitre. L'élément H^* de la figure 5.13 peut être envisagé comme un canal de transmission avec le bruit h^*, alors que les éléments A, G, B peuvent être réunis et nommés récepteur. Ce dernier problème diffère pourtant quelque peu des problèmes résolus au § 2. Citons certaines de ses particularités :

a) Le « récepteur» du système comprend les éléments donnés G et B. Le problème consiste à rechercher l'algorithme d'une seule

partie A de ce système. Constatons que la théorie des communications compte également des problèmes analogues.

b) A l'intérieur de l'organe « récepteur » du système il existe des perturbations aléatoires g_s et z_s. Un problème analogue peut également être posé pour un système de communication qui donne lieu à des bruits de fond.

c) Dès le début du processus il faut obtenir à la sortie x_s de l'objet B une grandeur telle qu'elle minimise l'espérance mathématique de la fonction de pertes. Ainsi, lorsque

$$W_s = (x_s^* - x_s)^2, \tag{5.95}$$

il faut, évidemment, minimiser une certaine mesure de l'écart de la « décision » x_s du « signal transmis » x_s^*. Dans le cas idéal $x_s = x_s^*$ et W_s atteint sa valeur minimale égale à zéro. Pourtant, les perturbations aléatoires qui se superposent sur le signal « reçu » y_s^* ainsi que les bruits de fond g_s et z_s du « récepteur » du système font que les erreurs de décision sont inévitables et, en général, x_s ne coïncide pas avec x_s^*. Cette formulation du problème diffère de celle du problème du § 2 par ce que dans ce dernier la décision ne devait être prise qu'à la fin de l'observation du signal reçu, alors que dans le cas considéré, la décision doit être élaborée continûment, simultanément avec la réception du signal x_s^* (décision courante).

Récrivons l'expression (5.94) sous la forme suivante :

$$R_s = \int\limits_{\Omega(x_s,\, v_s,\, u_s,\, \vec{y}_{s-1}^*)} P(x_s \mid v_s)\, P(v_s \mid u_s)\, \Gamma_s(u_s \mid \vec{y}_{s-1}^*) \times$$
$$\times \left\{ \int\limits_{\Omega(\bar{\lambda})} W_s[s, x_s^*, (s, \bar{\lambda}), x_s]\, P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*) \times \right.$$
$$\times P(\bar{\lambda})\, d\Omega(\bar{\lambda})\, d\Omega(x_s, u_s, v_s, \vec{y}_{s-1}^*). \tag{5.96}$$

Introduisons la fonction

$$\rho_s = \rho_s(x_s, \vec{y}_{s-1}^*) = \int\limits_{\Omega(\bar{\lambda})} W_s[s, x_s^*(s, \bar{\lambda}), x_s] \times$$
$$\times P(\vec{y}_{s-1}^* \mid \vec{x}_{s-1}^*)\, P(\bar{\lambda})\, d\Omega(\bar{\lambda}). \tag{5.97}$$

L'expression (5.96) peut se mettre alors sous une forme plus condensée :

$$R_s = \int\limits_{\Omega(x_s,\, v_s,\, u_s,\, \vec{y}_{s-1}^*)} P(x_s \mid v_s)\, P(v_s \mid u_s) \times$$
$$\times \Gamma_s(u_s \mid \vec{y}_{s-1}^*)\, \rho(x_s, \vec{y}_{s-1}^*)\, d\Omega. \tag{5.98}$$

La formule (5.98) montre que le choix de la fonction Γ_s avec s fixé n'influe que sur le terme R_s du risque total R. Par conséquent, il suffit de prendre dans la formule (5.98) une fonction Γ_s telle qu'elle minimise le risque élémentaire R_s. Cette fonction sera également optimale dans le sens de minimisation du risque total R. Γ_s étant la densité de probabilité, $\Gamma_s > 0$ et, de plus,

$$\int\limits_{\Omega(u_s)} \Gamma_s(u_s \mid \vec{y}^*_{s-1})\, d\Omega = 1. \tag{5.99}$$

L'expression (5.98) peut s'écrire :

$$R_s = \int\limits_{\Omega(u_s,\, \vec{y}^*_{s-1})} \Gamma_s(u_s \mid \vec{y}^*_{s-1}) \left\{ \int\limits_{\Omega(x_s,\, v_s)} P(x_s \mid v_s)\, P(v_s \mid u_s) \times \right.$$
$$\left. \times \rho_s(x_s, \vec{y}^*_{s-1})\, d\Omega(x_s, v_s) \right\} d\Omega(u_s, \vec{y}^*_{s-1}). \tag{5.100}$$

L'intégrale dans l'accolade est la fonction de u_s et de $\vec{y}^*_{s-1}$ que nous allons noter $\xi_s(u_s, \vec{y}^*_{s-1})$:

$$\xi_s(u_s, \vec{y}^*_{s-1}) = \int\limits_{\Omega(x_s,\, v_s)} P(x_s \mid v_s)\, P(v_s \mid u_s)\, \rho_s(x_s, \vec{y}^*_{s-1})\, d\Omega. \tag{5.101}$$

Cette fonction peut s'obtenir à partir des données du problème. (5.101) et (5.100) entraînent

$$R_s = \int\limits_{\Omega(u_s,\, \vec{y}^*_{s-1})} \xi_s(u_s, \vec{y}^*_{s-1})\, \Gamma_s(u_s \mid \vec{y}^*_{s-1})\, d\Omega =$$
$$= \int\limits_{\Omega(\vec{y}^*_{s-1})} I(\vec{y}^*_{s-1})\, d\Omega(\vec{y}^*_{s-1}), \tag{5.102}$$

où

$$I(\vec{y}^*_{s-1}) = \int\limits_{\Omega(u_s)} \Gamma_s(u_s \mid \vec{y}^*_{s-1})\, \xi_s(u_s, \vec{y}^*_{s-1})\, d\Omega. \tag{5.103}$$

La formule (5.102) montre que le risque R_s est minimal si pour tout $\vec{y}^*_{s-1}$ la fonction I est minimale. Par conséquent, dans l'expression (5.103) il faut choisir Γ_s tel qu'il minimise l'intégrale $I(\vec{y}^*_{s-1})$ pour une valeur quelconque du paramètre $\vec{y}^*_{s-1}$.

Le théorème de la moyenne et le fait que les fonctions sous le signe somme de (5.103) sont positives permettent d'écrire

$$I(\vec{y}^*_{s-1}) = (\xi_s)_{\text{moy}} \int\limits_{\Omega(u_s)} \Gamma_s(u_s \mid \vec{y}^*_{s-1})\, d\Omega = (\xi_s)_{\text{moy}} \geqslant (\xi_s)_{\min}, \tag{5.104}$$

où $(\xi_s)_{\text{moy}}$ est la valeur moyenne et $(\xi_s)_{\min}$ la valeur minimale de ξ_s. D'après (5.99) l'intégrale dans (5.104) est égale à l'unité. Par suite de l'expression (5.104), la valeur minimale possible de $I(\vec{y}^*_{s-1})$ est égale à $(\xi_s)_{\min}$. On montre sans peine que cette valeur de I s'obtient si la fonction Γ_s est choisie de la façon suivante *). Soit u_s^* la valeur de u_s qui minimise la fonction $\xi_s(u_s)$ dans le domaine $\Omega(u_s)$ des valeurs possibles de u_s. De plus, en donnant des conditions supplémentaires on peut limiter d'une certaine façon le domaine $\Omega(u_s)$. Il se peut que $\xi_s(u_s^*)$ est le plus petit de plusieurs minimums locaux de la fonction $\xi_s(u_s)$. Quoi qu'il en soit, considérons que nous avons trouvé une valeur de u_s telle que

$$\xi_s(u_s^*) = \min_{(u_s)} \xi_s(u_s). \tag{5.105}$$

Considérons maintenant la fonction

$$\Gamma_s(u_s) = \delta(u_s - u_s^*), \tag{5.106}$$

où δ est l'impulsion unité (impulsion de Dirac). Il est clair que la fonction Γ_s^* vérifie la condition (5.99), car l'intégrale de l'impulsion unité sur tout le domaine $\Omega(u_s)$ est égale à l'unité. Il se trouve que l'expression (5.106) donne la fonction optimale Γ_s^*. Pour le montrer, signalons une propriété connue de la fonction δ : pour toute fonction continue $\varphi(x)$

$$\int_{-\infty}^{\infty} \delta(x - x^*)\, \varphi(x)\, dx = \varphi(x^*). \tag{5.107}$$

Cette propriété se généralise aisément au cas de l'intégration multiple, c'est-à-dire de l'intégration sur un domaine multidimensionnel.

Portons (5.106) dans le premier membre de (5.103). Il vient

$$[I(\vec{y}^*_{s-1})]_{\Gamma_s=\delta(u_s-u_s^*)} = \int_{\Omega(u_s)} \delta(u_s - u_s^*)\, \xi_s(u_s, \vec{y}^*_{s-1})\, d\Omega =$$

$$= \xi_s(u_s^*) = (\xi_s)_{\min} = (I)_{\min}. \tag{5.108}$$

Par suite, I atteint sa valeur minimale lors de l'utilisation d'une stratégie (5.106).

Ainsi, comme on le voit de la formule (5.106), la stratégie optimale Γ_s^* est une stratégie pure. Dans le cas considéré, on pouvait d'ailleurs le prévoir puisqu'il s'agit d'un problème de Bayes et les probabilités a priori sont données. Nous avons montré au paragraphe

---

*) $\vec{y}^*_{s-1}$ est considéré ici comme un paramètre et ne s'écrit pas sous une forme explicite.

précédent que dans la théorie des décisions bialternatives la stratégie optimale de Bayes est une stratégie pure. Cette règle se confirme également pour un cas plus complexe.

L'algorithme optimal d'un organe de commande consiste à choisir la valeur u_s^* telle qu'elle minimise la fonction ξ_s

$$\xi_s(u_s^*, \vec{y}_{s-1}^*) = \min_{(u_s)} \xi_s(u_s, \vec{y}_{s-1}^*). \tag{5.109}$$

On en déduit que u_s^* est une fonction de $\vec{y}_{s-1}^*$:

$$u_s^* = u_s^*(\vec{y}_{s-1}^*). \tag{5.110}$$

Ainsi, comme il fallait s'y attendre, le choix optimal de la valeur de u_s à l'instant $t = s$ dépend généralement de toute la « préhistoire » des signaux y_i^* observés par l'organe A, c'est-à-dire de la suite des valeurs $y_0^*, y_1^* \ldots, y_{s-1}^*$.

Si la fonction $\xi_s(u_s, \vec{y}_{s-1}^*)$ est calculée au préalable, en y portant le vecteur $\vec{y}_{s-1}^*$ enregistré par l'organe de commande A et en minimisant ξ_s par rapport à u_s on peut trouver la valeur recherchée

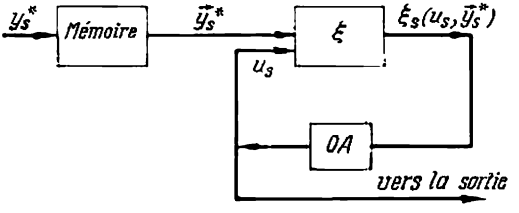


Fig. 5.15.

de la commande optimale courante u_s^*. Le schéma fonctionnel d'une version possible de l'organe de commande optimal A est représenté sur la figure 5.15. Les valeurs courantes y_s^* sont fournies à la mémoire où s'accumule la suite $y_0^*, \ldots, y_{s-1}^*$, c'est-à-dire le vecteur $\vec{y}_{s-1}^*$. La valeur de ce vecteur est ensuite transmise au bloc ξ dans lequel s'élabore la fonction $\xi_s(u_s, \vec{y}_{s-1}^*)$. La valeur de u_s de la sortie de l'optimaliseur automatique OA est appliquée à la deuxième entrée du bloc ξ. L'optimaliseur automatique choisit la valeur de $u_s = u_s^*$. L'optimaliseur automatique choisit la valeur de $u_s = u_s^*$ de façon qu'elle minimise la fonction ξ_s. La valeur u_s^* apparaît à la sortie du bloc A. Bien entendu, dans les schémas réels nul besoin n'est de retenir chacune des valeurs $y_0^*, y_1^*, \ldots, y_{s-1}^*$. Il ne faut garder en mémoire qu'un petit nombre de fonctions $\psi_{s1}, \ldots, \psi_{sp}$

de ces valeurs où p n'est pas grand, dites statistiques exhaustives (coordonnées suffisantes).

Il est difficile souvent d'obtenir la formule de $\xi_s(u_s, \vec{y}^*_{s-1})$ sous une forme explicite. Dans ce cas, le bloc ξ peut représenter un calculateur qui effectue automatiquement l'intégration d'après l'expression (5.101). A la sortie de ce bloc on obtient la valeur numérique de ξ_s qui correspond aux valeurs de $\vec{y}^*_{s-1}$ et u_s à ses entrées.

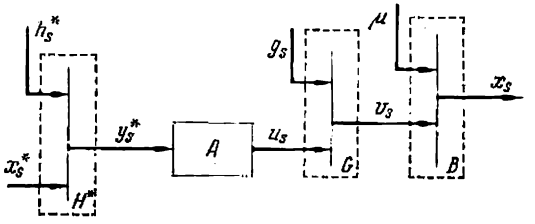


Fig. 5.16.

Considérons à titre d'exemple le schéma fonctionnel d'un système simple représenté sur la figure 5.16.

Les équations de ce système s'écrivent

$$\left.\begin{aligned} y_s^* &= x_s^* + h_s^*, \\ v_s &= u_s + g_s, \\ x_s &= v_s + \mu. \end{aligned}\right\} \tag{5.111}$$

La première de ces équations décrit les propriétés de la chaîne H^*, la deuxième de la chaîne G. La dernière équation est celle de l'objet B.

Soient h_s^* et g_s les suites des variables aléatoires indépendantes normales, de moyennes nulles

$$\left.\begin{aligned} P(h_s^*) &= \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{(h_s^*)^2}{2\sigma_h^2}\right\}, \\ P(g_s) &= \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{(g_s)^2}{2\sigma_g^2}\right\}. \end{aligned}\right\}$$

Dans l'exemple considéré le vecteur μ est remplacé par la variable aléatoire μ de densité de probabilité

$$P(\mu) = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{\mu^2}{2\sigma_\mu^2}\right\}. \tag{5.112}$$

Soit $x_s^* = \lambda$, où λ est la variable aléatoire normale de densité de probabilité a priori

$$P(\lambda) = \frac{1}{\sigma_\lambda \sqrt{2\pi}} \exp\left\{-\frac{(\lambda-\lambda_0)^2}{2\sigma_\lambda^2}\right\}. \tag{5.113}$$

On se donne la fonction de pertes sous la forme

$$W = \sum_{s=0}^{n} W_s = \sum_{s=0}^{n} (x_s^* - x_s)^2. \tag{5.114}$$

Il faut chercher l'algorithme optimal de l'organe de commande A tel qu'il minimise l'espérance mathématique de la fonction W.

Cherchons d'abord la fonction ρ_s en utilisant l'expression (5.97) :

$$\rho_s = \rho_s(x_s, \vec{y}_{s-1}^*) = \int_{-\infty}^{\infty} (\lambda - x_s)^2 \prod_{i=0}^{s-1} P(y_i^* \mid \lambda) P(\lambda)\, d\lambda. \tag{5.115}$$

Etant donné que $h_i^* = y_i^* - x_i^* = y_i^* - \lambda$,

$$P(y_i^* \mid \lambda) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{(y_i^* - \lambda)^2}{2\sigma_h^2}\right\}. \tag{5.116}$$

En portant (5.116) et (5.113) dans l'expression (5.115), il vient :

$$\rho_s = \frac{1}{(\sigma_h)^s \sigma_\lambda (2\pi)^{\frac{s+1}{2}}} \times$$

$$\times \int_{-\infty}^{\infty} (\lambda - x_s)^2 \exp\left\{-\frac{\sum_{i=0}^{s-1} (y_i - \lambda)^2}{2\sigma_h^2} - \frac{(\lambda - \lambda_0)^2}{2\sigma_\lambda^2}\right\} d\lambda. \tag{5.117}$$

Introduisons les notations

$$\left.\begin{aligned} a &= \frac{1}{(\sigma_h)^s \sigma_\lambda (2\pi)^{\frac{s+1}{2}}}, \quad A_s = \frac{\sigma_h^2 + s\sigma_\lambda^2}{2\sigma_h^2\sigma_\lambda^2}, \\ B_{s-1} &= \frac{\sigma_\lambda^2 \sum_{i=0}^{s-1} y_i^* + \lambda_0 \sigma_h^2}{\sigma_h^2 \sigma_h^2}, \quad C_{s-1} = \frac{\sigma_\lambda^2 \sum_{i=0}^{s-1} (y_i^*)^2 + \lambda_0 \sigma_h^2}{2\sigma_\lambda^2 \sigma_\lambda^2} \end{aligned}\right\} \tag{5.118}$$

La formule (5.117) s'écrit alors

$$\rho_s = a \exp\{-C_{s-1}\} \int_{-\infty}^{\infty} (\lambda - x_s)^2 \exp\{-A_s\lambda^2 + B_{s-1}\lambda\}\, d\lambda. \tag{5.119}$$

Utilisons la formule connue (cf. par exemple, [5.26])

$$\int_{-\infty}^{\infty} x^2 \exp\{-px^2 + 2qx\}\, dx = \frac{1}{2p}\sqrt{\frac{\pi}{p}}\left(1 + \frac{2q^2}{p}\right) \exp\left\{\frac{q^2}{p}\right\}. \tag{5.120}$$

Appliquons cette formule à (5.119) et posons $z = \lambda - x_s$. Avec cette substitution l'intégrale (5.119) devient

$$\rho_s = \frac{a\sqrt{\pi}}{2A_s^{3/2}}\left[1 + \frac{2}{A_s}\left(A_s x_s - \frac{B_{s-1}}{2}\right)^2\right] \times$$

$$\times \exp\left\{-C_{s-1} + \frac{B_{s-1}}{4A_s}\right\}. \qquad (5.121)$$

Cherchons maintenant $\xi_s(u_s, y^*_{s-1})$. Puisque $\mu = x_s - v_s$, on a

$$P(x_s \mid v_s) = \frac{1}{\sigma_\mu\sqrt{2\pi}} \exp\left\{-\frac{(x_s - v_s)^2}{2\sigma_\mu^2}\right\}. \qquad (5.122)$$

Ensuite, étant donné que $g_s = v_s - u_s$, on a

$$P(v_s \mid u_s) = P(g_s = v_s - u_s) = \frac{1}{\sigma_g\sqrt{2\pi}} \exp\left\{-\frac{(v_s - u_s)^2}{2\sigma_g^2}\right\}. \qquad (5.123)$$

D'après (5.101) et (5.121) et compte tenu de (5.122) et de (5.123) on obtient

$$\xi_s = b \exp\left\{-C_{s-1} + \frac{B_{s-1}^2}{4A_s}\right\} \int\limits_{x_s=-\infty}^{\infty} \left[1 + \frac{2}{A_s}\left(A_s x_s - \frac{B_{s-1}}{2}\right)^2\right] \times$$

$$\times \left\{\int\limits_{-\infty}^{\infty} \exp\left\{-\frac{(x_s - v_s)^2}{2\sigma_\mu^2} - \frac{(v_s - u_s)^2}{2\sigma_g^2}\right\} dv_s\right\} dx_s, \qquad (5.124)$$

où

$$b = \frac{a\sqrt{\pi}}{2A_s^{3/2}} \cdot \frac{1}{\sigma_\mu \sigma_g 2\pi}. \qquad (5.125)$$

Posons

$$\left.\begin{aligned} \alpha_s &= \frac{\sigma_g^2 x_s^2 + \sigma_\mu^2 u_s^2}{2\sigma_\mu^2 \sigma_g^2}, \\ \beta_s &= \frac{\sigma_g^2 x_s + \sigma_\mu^2 u_s}{\sigma_\mu^2 \sigma_g^2}. \\ \gamma &= \frac{\sigma_g^2 + \sigma_\mu^2}{2\sigma_\mu^2 + \sigma_g^2}. \end{aligned}\right\} \qquad (5.126)$$

L'intégrale dans l'accolade de (5.124) devient

$$I_0 = \exp\{-\alpha_s\} \int\limits_{-\infty}^{\infty} \exp\{-\gamma v_s^2 + \beta_s v_s\}\, dv_s. \qquad (5.127)$$

En appliquant la formule (cf. [5.26])

$$\int\limits_{-\infty}^{\infty} \exp\{-px^2 \pm qx\}\, dx = \sqrt{\frac{\pi}{p}} \exp\left\{\frac{q^2}{4p}\right\}, \qquad (5.128)$$

on aboutit après calcul à

$$I_0 = \exp\left\{-\alpha_s + \frac{\beta_s^2}{4\gamma}\right\}\sqrt{\frac{\pi}{\gamma}}. \tag{5.129}$$

En portant (5.128) dans (5.124), on obtient l'expression

$$\xi_s = \left[b\sqrt{\frac{\pi}{\gamma}}\exp\left\{-C_{s-1} + \frac{B_{s-1}^2}{4A_s}\right\}\right]\left(I_1 + \frac{2}{A_s}I_2\right), \tag{5.130}$$

où I_1 et I_2 sont les intégrales écrites ci-dessous. Posons

$$\left.\begin{aligned} L &= \frac{1}{2\sigma_\mu^2} - \frac{1}{4\gamma\sigma_\mu^4}, \\ M &= \frac{1}{2\gamma\sigma_\mu^2\sigma_g^2}, \quad N = -\frac{1}{2\sigma_g^2} + \frac{1}{4\gamma\sigma_g^4}. \end{aligned}\right\} \tag{5.131}$$

Alors

$$I_1 = \sqrt{\frac{\pi}{L}}\exp\left\{\left(N + \frac{M^2}{4L}\right)u_s^2\right\}. \tag{5.132}$$

Pour obtenir cette expression on a appliqué la formule (5.128). La formule (5.120) conduit ensuite à

$$I_2 = \int_{-\infty}^{\infty}\left(A_s x_s - \frac{\beta_{s-1}}{2}\right)^2 \exp\{-Lx_s^2 + Mx_s u_s + Nu_s^2\}\,dx_s =$$

$$= \frac{A_s^2}{2L}\sqrt{\frac{\pi A_s^2}{L}}\left[1 + \frac{A_s^2}{2L}\left(-\frac{LB_{s-1}}{A_s^2} + \frac{Mu_s}{A_s}\right)^2\right]\exp\left\{\left(N + \frac{M^2}{4L}\right)u_s^2\right\}. \tag{5.133}$$

En portant (5.132) et (5.133) dans (5.130) on obtient après des transformations évidentes

$$\xi_s = D_s[1 + \varepsilon^2(\vartheta u_s - 1)^2], \tag{5.134}$$

de plus il se trouve que

$$N + \frac{M^2}{4L} = 0.$$

Ici

$$\left.\begin{aligned} D_s &= \left(1 + \frac{A_s^2}{L}\right)\sqrt{\frac{\pi}{L}}\,b\sqrt{\frac{\pi}{\gamma}}\exp\left\{-C_{s-1} + \frac{B_{s-1}^2}{4A_s}\right\}, \\ \varepsilon^2 &= \frac{LB_{s-1}^2}{2(L + A_s^2)}, \quad \vartheta = \frac{2A_s}{B_{s-1}}. \end{aligned}\right\} \tag{5.135}$$

Les expressions D_s et ε^2 ne dépendent pas de u_s. C'est pourquoi le minimum de ξ_s par rapport à u_s s'obtient si l'on annule l'expression entre parenthèses de (5.134) :

$$\vartheta u_s - 1 = 0. \tag{5.136}$$

On en tire la commande optimale u_s^* sous la forme

$$u_s^* = \frac{1}{\vartheta} = \frac{\lambda_0 + \left(\frac{\sigma_\lambda}{\sigma_h}\right)^2 \sum_{i=1}^{s-1} y_i^*}{1 + s\left(\frac{\sigma_\lambda}{\sigma_h}\right)^2}. \tag{5.137}$$

Le sens de cette formule est qu'elle donne l'estimation de la valeur λ. La commande u_s doit reproduire de la meilleure façon la valeur λ; la valeur x_s répondra alors le mieux à la grandeur $x_s^* = \lambda$.

Comme on le voit de l'expression (5.137), lorsque les valeurs de s et de $\left(\frac{\sigma_\lambda}{\sigma_h}\right)^2$ sont petites (par exemple, dans le cas d'une variance σ_h^2 importante du bruit), l'estimation est basée sur la valeur moyenne a priori λ_0. Toutefois, lorsque s est suffisamment grand, la valeur λ_0 ne joue plus un rôle important du fait que pendant le temps s s'accumule la somme $\sum_{i=0}^{s-1} y_i^*$ et en pratique sa valeur dépasse nettement λ_0. Dans ce cas la formule (5.137) donne

$$u_s^* \cong \frac{\sum_{i=0}^{s-1} y_i^*}{s}, \tag{5.138}$$

c'est-à-dire l'estimation de la grandeur λ se ramène à l'obtention de la moyenne arithmétique des valeurs y_i^* $(i = 0, 1, \ldots, s-1)$, mesurées à l'entrée de l'organe de commande.

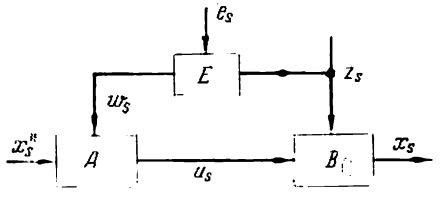


Fig. 5.17.

D'une manière analogue on peut construire une théorie de compensation optimale. Examinons, par exemple, le système dont le schéma fonctionnel est représenté sur la figure 5.17.

Supposons, pour simplifier l'exposé, que le bruit aléatoire n'existe que dans le circuit E de mesure de la perturbation z_s. Supposons encore que x_s^* est donné d'avance. Soit le système B sans mémoire décrit par l'équation

$$x_s = F(z_s, u_s),$$

où, comme auparavant,

$$z_s = z_s(s, \bar{\mu}),$$

et la densité de probabilité a priori $P(\bar{\mu})$ du vecteur $\bar{\mu}$ est donnée. La densité de probabilité $P(e_s)$ du bruit e_s dans la chaîne E de mesure de la perturbation z_s est également donnée. Supposons que le bruit e_s est une suite de variables indépendantes, suivant une même loi de répartition $P(e_s)$.

Introduisons les vecteurs temporels

$$\left.\begin{aligned} \vec{z}_k &= (z_0,\ z_1,\ \ldots,\ z_k), \\ \vec{w}_k &= (w_0,\ w_1,\ \ldots,\ w_k), \\ \vec{x}_k^* &= (x_0^*,\ x_1^*,\ \ldots,\ x_k^*). \end{aligned}\right\}$$

Le problème consiste à déterminer les stratégies optimales $\Gamma_s(u_s \mid \vec{x}_s^*,\ \vec{w}_{s-1})$ de l'organe de commande A $(s = 0,\ 1,\ \ldots,\ n)$. Les expressions de la fonction de pertes W et du risque R restent les mêmes que dans la théorie exposée plus haut dans ce paragraphe.

Déduisons d'abord la formule du risque conditionnel élémentaire r_s, le vecteur $\vec{z}_s$ étant fixé :

$$r_s = M\{W_s \mid \vec{z}_s\} = \int\limits_{\Omega(u_s)} W_s[s,\ x_s^*,\ F(z_s,\ u_s)]\, P_s(u_s \mid \vec{x}_s^*,\ \vec{z}_{s-1})\, d\Omega. \quad (5.139)$$

Dans cette formule x_s est remplacé dans l'expression de W_s par son expression. C'est pourquoi dans l'expression de W_s l'unique grandeur aléatoire est u_s et l'espérance mathématique $M\{W_s\}$ s'obtient en prenant la moyenne par rapport à u_s. Soit $P_s(u_s \mid \vec{x}_s^*,\ \vec{z}_{s-1})$ une densité de probabilité conditionnelle de u_s pour $\vec{z}_{s-1}$ et $\vec{x}_s^*$ fixés.

La fonction $P(u_s \mid \vec{x}_s^*,\ \vec{z}_{s-1})$ est fournie par la formule qui se déduit d'une façon évidente du schéma fonctionnel représenté sur la figure 5.17 :

$$P_s(u_s \mid \vec{x}_s,\ \vec{z}_{s-1}) = \int\limits_{\Omega(\vec{w}_{s-1})} \Gamma_s(u_s \mid \vec{x}_s^*,\ \vec{w}_{s-1})\, P(\vec{w}_{s-1} \mid \vec{z}_{s-1})\, d\Omega. \quad (5.140)$$

Γ_s est ici la stratégie de l'organe de commande et $P(\vec{w}_{s-1} \mid \vec{z}_{s-1})$ la densité de probabilité conditionnelle du vecteur de sortie de l'élément E, le vecteur $\vec{z}_{s-1}$ de la perturbation étant fixé. Puisque w_i sont des variables indépendantes, on obtient

$$P(\vec{w}_{s-1} \mid \vec{z}_{s-1}) = \prod_{i=0}^{s-1} P(w_i \mid z_i). \quad (5.141)$$

En portant (5.141) dans (5.140) et (5.140) dans (5.139) on a

$$r_s = \int\limits_{\Omega(u_s,\, \vec{w}_{s-1})} W_s\{s,\, x_s^*,\, F[z_s(s,\, \bar{\mu}),\, u_s]\}\, \Gamma_s(u_s \mid \vec{x}_s^*,\, \vec{w}_{s-1}) \times$$

$$\times \prod_{i=0}^{s-1} \left\{ P[w_i \mid z_i(i,\, \bar{\mu})] \right\} d\Omega. \qquad (5.142)$$

Les vecteurs $\vec{z}_s$ varient suivant l'expérience. Si l'on connaît la densité de probabilité $P(\bar{\mu})$, on peut obtenir le risque élémentaire total R_s d'après la formule

$$R_s = \int\limits_{\Omega(\bar{\mu})} r_s(\bar{\mu})\, P(\bar{\mu})\, d\Omega.$$

En remplaçant ici r_s par l'expression (5.142), il vient :

$$R_s = \int\limits_{\Omega(u_s,\, \vec{w}_{s-1},\, \bar{\mu})} W_s\{s,\, x_s^*,\, F[z_s(s,\, \bar{\mu}),\, u_s]\}\, \Gamma_s(u_s \mid \vec{x}_s^*,\, \vec{w}_{s-1}) \times$$

$$\times \left\{ \prod_{i=0}^{s-1} P[w_i \mid z_i(i,\, \bar{\mu})] \right\} P(\bar{\mu})\, d\Omega. \qquad (5.143)$$

Le risque R_s ne dépendant que de la stratégie Γ_s associée à l'instant considéré $t = s$, la stratégie optimale Γ_s^* est celle qui minimise le risque élémentaire R_s. Posons

$$\eta_s(u_s,\, x_s^*,\, \vec{w}_{s-1}) = \int\limits_{\Omega(\bar{\mu})} W_s\{s,\, x_s^*,\, F[z_s(s,\, \bar{\mu}),\, u_s]\} \times$$

$$\times \left\{ \prod_{i=0}^{s-1} P[w_i \mid z_i(i,\, \bar{\mu})] \right\} P(\bar{\mu})\, d\Omega. \qquad (5.144)$$

On a

$$R_s = \int\limits_{\Omega(\vec{w}_{s-1},\, u_s)} \Gamma_s(u_s \mid \vec{x}_s^*,\, \vec{w}_{s-1})\, \eta_s(u_s,\, x_s^*,\, \vec{w}_{s-1})\, d\Omega =$$

$$= \int\limits_{\Omega(\vec{w}_{s-1})} I(\vec{w}_{s-1})\, d\Omega, \qquad (5.145)$$

où

$$I(\vec{w}_{s-1}) = \int\limits_{\Omega(u_s)} \Gamma_s(u_s \mid \vec{x}_s^*,\, w_{s-1})\, \eta_s(u_s,\, x_s^*,\, \vec{w}_{s-1})\, d\Omega. \qquad (5.146)$$

Si l'on choisit Γ_s telle qu'elle minimise la grandeur $I(\vec{w}_{s-1})$ pour tout $\vec{w}_{s-1}$, la valeur de R_s sera alors également minimale. Cette sélection s'effectue de la même façon que dans la théorie exposée dans ce qui précède.

Soit

$$\eta_s(u_s^*, x_s^*, \vec{w}_{s-1}) = \min_{(u_s)} \eta_s(u_s, x_s^*, \vec{w}_{s-1}). \tag{5.147}$$

Cherchons la valeur u_s^* qui minimise η_s. Il est évident que u_s^* est en général fonction de x_s^* et de $\vec{w}_{s-1}$:

$$u_s^* = u_s^*(x_s^*, \vec{w}_{s-1}). \tag{5.148}$$

Or, c'est précisément l'expression de la commande optimale qui dépend de la valeur courante x_s et de toute la « préhistoire » des

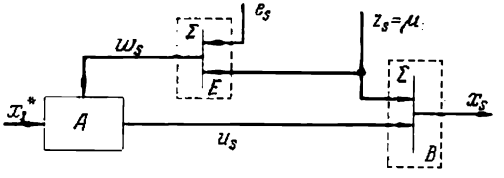


Fig. 5.18.

grandeurs d'entrée w_i $(i = 0, 1, \ldots, s-1)$, enregistrées par l'organe de commande A aux instants $t = 0, 1, \ldots, s-1$.

La stratégie optimale Γ_s^* est pure:

$$\Gamma_s^* = \delta(u_s - u_s^*),$$

où u_s^* est définie par la condition (5.147). La démonstration de la validité de cette expression est la même que dans la théorie exposée plus haut.

Considérons un exemple très simple du schéma donné par la figure 5.18. Soit $z_s = \mu$, tandis que les éléments E et B sont décrits par les équations

$$w_s = e_s + \mu, \quad x_s = \mu + u_s.$$

Soient ensuite les variables aléatoires normales e_s et μ caractérisées par les densités de probabilité

$$P(\mu) = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right\},$$

$$P(e_s) = \frac{1}{\sigma_e \sqrt{2\pi}} \exp\left\{-\frac{e_s^2}{2\sigma_e^2}\right\}.$$

Adoptons la fonction de pertes élémentaire sous la forme

$$W_s = (x_s^* - x_s)^2 = (x_s - \mu - u_s)^2. \tag{5.149}$$

Etant donné que $e_i = w_i - \mu$, il vient :

$$P = (w_i \mid z_s) = \frac{1}{\sigma_e \sqrt{2\pi}} \exp\left\{-\frac{(w_i - \mu)^2}{2\sigma_e}\right\}. \tag{5.150}$$

Cherchons d'après la formule (5.144) l'expression de η_s :

$$\eta_s = a \int_{-\infty}^{\infty} (x_s^* - \mu - u_s)^2 \exp\left\{-\frac{\sum_{i=0}^{s-1} (w_i - \mu)^2}{2\sigma_e^2} - \frac{(\mu - \mu_0)^2}{2\sigma_\mu^2}\right\} d\mu, \tag{5.151}$$

où a est une constante. Transformons l'expression dans l'accolade sous le signe somme en effectuant le remplacement $\alpha = \mu + u_s - x_s^* = \mu - y_s$, où $y_s = x_s^* - u_s$.

Après les transformations, l'expression (5.151) devient

$$\eta_s = a \exp\{-A_s\} \int_{-\infty}^{\infty} \alpha^2 \exp\{-C_s\alpha^2 - B_s\alpha\}\, d\alpha, \tag{5.152}$$

où

$$\left.\begin{aligned} A_s &= \frac{\left(\sigma_\mu^2 \sum_{i=0}^{s-1} w_i^2 + \sigma_e^2\mu_0\right) - y_s\left(2\sigma_\mu^2 \sum_{i=0}^{s-1} w_i + 2\sigma_e^2\mu_0\right) + (s\sigma_\mu^2 + \sigma_e^2)\, y_s^2}{2\sigma_e^2\sigma_\mu^2}. \\ B_s &= \frac{-\sigma_\mu^2 \sum_{i=0}^{s-1} w_i - \sigma_e^2\mu_0 + y_s(s\sigma_\mu^2 + \sigma_e^2)}{2\sigma_e^2\sigma_\mu^2}, \\ C_s &= \frac{s\sigma_\mu^2 + \sigma_e^2}{2\sigma_e^2\sigma_\mu^2}. \end{aligned}\right\} \tag{5.153}$$

L'intégrale (5.152) est définie par la formule (5.120). En définitive on a

$$\eta_s = b \exp\left\{-A_s + \frac{B_s^2}{4C_s}\right\} \cdot \left[1 + \frac{B_s^2}{4C_s}\right], \tag{5.154}$$

où b est la constante comportant également certains facteurs dépendant de s. En posant

$$\beta_s = \frac{\sigma_\mu \sum_{i=0}^{s-1} w_i + \sigma_e^2\mu_0}{\sigma_e\sigma_\mu^2}$$

et en comparant les deux dernières formules dans (5.153) on obtient

$$B_s = 2y_s C_s - \beta_s. \tag{5.155}$$

En considérant l'exposant dans (5.154), on remarque qu'il ne contient pas y_s et, par suite, ne dépend pas de u_s. En effet, (5.155) entraîne

$$\frac{B_s}{4C_s} = \frac{(2y_s C_s - \beta_s)^2}{4C_s} = y_s^2 C_s - y_s \beta_s + \frac{\beta_s^2}{4C_s}. \tag{5.156}$$

Comparons l'expression (5.156) avec la formule de A_s dans (5.153). Dans ces expressions les termes qui comportent y_s^2 sont identiques. Il en est de même des termes qui comportent y_s. Par conséquent, ces termes disparaissent de l'exposant et la différence $\{-A_s + B_s^2/4C_s\}$ ne dépend pas de y_s et ni donc de u_s. C'est pourquoi minimiser η_s par rapport à u_s revient à minimiser le facteur $[1 + B_s^2/2C_s]$, où B_s dépend de u_s. Ce facteur est minimal quand $B_s = 0$. D'après l'expression (5.155), cette condition est vérifiée par la valeur

$$y_s = \frac{\beta_s}{2C_s}. \tag{5.157}$$

En remplaçant y_s et β_s par leurs valeurs, on obtient la loi de commande optimale

$$u_s^* = x_s^* - \frac{\mu_0 + \left(\frac{\sigma_\mu}{\sigma_e}\right)^2 \sum_{i=0}^{s-1} w_i}{1 + \left(\frac{\sigma_\mu}{\sigma_e}\right)^2 s}. \tag{5.158}$$

Le deuxième terme de cette formule est une estimation de la grandeur μ. En effet, la figure 5.17 montre que

$$u_s = x_s^* - \mu \tag{5.159}$$

serait la valeur idéale du fait que c'est précisément sous cette condition que x_s est égal à x_s^* et la fonction de pertes élémentaire W_s s'annule. En comparant (5.159) et (5.158) on voit aisément que le deuxième terme de (5.158) donne justement l'estimation de μ la plus avantageuse du point de vue du critère d'optimalité appliqué.

D'après (5.158), pour des s petits et des rapports $\frac{\sigma_\mu}{\sigma_e}$ faibles, par exemple, dans le cas d'une grande variance du bruit e_s, le rôle principal dans l'estimation de μ revient à la valeur moyenne a priori μ_0. Mais lorsque les s sont suffisamment grands, lorsque la somme $\sum_{i=0}^{s-1} w_i$ accumulée devient importante, la formule (5.158) peut s'écrire sous

une forme approchée

$$u_s^* \cong x_s^* - \frac{\sum_{i=0}^{s-1} w_i}{s} \tag{5.160}$$

et l'estimation de μ se ramène à une moyenne arithmétique des w_i.

Au § 1 de ce chapitre nous avons montré qu'il est possible en principe de remplacer la mesure directe de la perturbation z_s par une mesure indirecte de l'entrée et de la sortie de l'objet B. Considérons, par exemple, le schéma d'un système à réaction représenté sur la figure 5.19. Ce schéma ne prévoit pas de mesure directe de la perturbation $z_s = \mu$, mais la chaîne de réaction est munie d'un canal H dans lequel au signal x_s s'ajoute un bruit représentant une suite de variables aléatoires normales indépendantes.

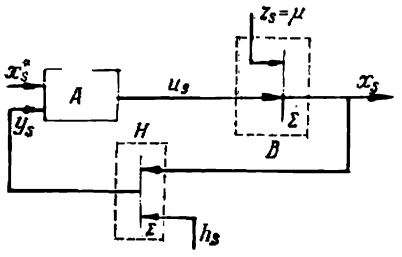


Fig. 5.19.

Soient les équations des éléments B et H

$$x_s = u_s + z_s = u_s + \mu,$$
$$y_s = x_s + h_s. \tag{5.161}$$

La grandeur de sortie y_s du canal H est appliquée à l'organe de commande A. (5.161) entraîne

$$y_s - u_s = \mu + h_s;$$

la différence $(y_s - u_s)$ fournie par l'organe de commande est en quelque sorte le résultat de mesure de la grandeur $z_s = \mu$ avec une erreur h_s. Ainsi, les schémas fonctionnels représentés sur les figures 5.17 et 5.18 sont parfaitement équivalents, à condition que

$$e_s = h_s.$$

En résumé, le schéma de la figure 5.18 se ramène à celui d'un système en boucle ouverte et, par conséquent, la théorie exposée ci-dessus, construite pour le système de la figure 5.17, convient très bien à l'étude du schéma de la figure 5.19. Pour abréger l'exposé nous appellerons *réductibles* des systèmes en boucle fermée qui peuvent être ramenés aux systèmes en boucle ouverte.

D'une manière analogue le système à réaction donné à titre d'exemple dans [5.25] peut être ramené à un système en boucle ouverte.

CHAPITRE VI

SYSTÈMES OPTIMAUX À ACCUMULATION ACTIVE DE L'INFORMATION

§ 1. Formulation d'un problème simple relatif à un système optimal de commande duale

La figure 6.1 représente le schéma fonctionnel d'un système de commande automatique en boucle fermée. La consigne $\bar{x}^*$ est amenée à l'entrée de l'organe de commande A par la chaîne H^* dans laquelle elle est entachée d'un bruit $\bar{h}^*$. C'est pourquoi l'action $\bar{y}^*$ appliquée

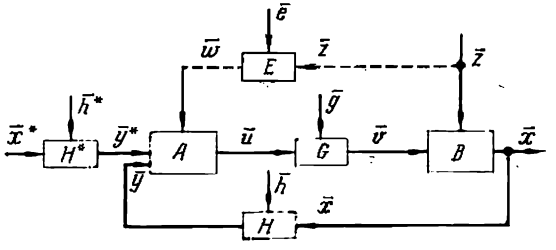


Fig. 6.1.

directement à l'entrée A n'est pas égale à la valeur réelle de la consigne $\bar{x}^*$. Il en est de même pour le signal $\bar{x}$ qui renseigne sur l'état du système gouverné B et qui est entaché dans la chaîne H du bruit $\bar{h}$. L'élément H fait partie de la chaîne de réaction ; sa grandeur de sortie $\bar{y}$ est amenée à l'entrée de l'organe de commande A. Ensuite la commande $\bar{u}$ passe à l'entrée du système B par la chaîne G où elle est entachée d'un bruit $\bar{g}$. C'est pourquoi l'action $\bar{v}$ introduite dans le système B n'est pas égale à $\bar{u}$. Le trait interrompu matérialise la boucle de compensation qui peut être adjointe au système. La valeur de la perturbation z qui agit sur le système B est mesurée pour être transmise à l'organe de commande A par le canal E où elle est entachée d'un bruit $\bar{e}$. Il en résulte que l'action $\bar{w}$ fournie à l'entrée de l'organe A n'est plus égale à la grandeur $\bar{z}$.

Dans le schéma en boucle fermée de la figure 6.1 sont possibles des processus qui n'ont pas d'analogue dans les systèmes en boucle ouverte étudiés au chapitre précédent. Un schéma en boucle ouverte

(fig. 5.12 ou 5.16, par exemple) ne permet qu'une observation passive des actions $\bar{x}^*$ ou $\bar{z}$. Or, dans le système en boucle fermée de la figure 6.1 l'étude de la perturbation $\bar{z}$ peut être active. Pour mieux explorer le système B et obtenir une information plus complète sur la perturbation $\bar{z}$ ou plutôt sur les variations aléatoires des caractéristiques du système, ce dernier peut être sondé par des tests réalisés dans le but de l'étudier.

Pourtant les actions de commande sont nécessaires non seulement pour l'étude du système, mais encore pour le ramener au régime requis. C'est pourquoi dans le schéma de la figure 6.1 les commandes doivent avoir un caractère *dual*, leur but devant être l'étude mais aussi la commande.

Le système dans lequel les actions de commande ont le caractère dual s'appelle système de commande duale [6.1]. Ce chapitre traite de la théorie de tels systèmes.

Une commande duale est avantageuse et même indispensable lorsque l'opérateur F du système B et la perturbation $\bar{z}$ ont une allure complexe rendant ainsi complexes et variables les caractéristiques du système. Les exemples des systèmes de commande duale sont fournis par les systèmes de recherche automatique et, notamment, les systèmes d'optimisation automatique [3.25, 6.2-6.6]. Dans les systèmes usuels de ce type la partie « exploratrice » de l'action peut facilement être séparée de sa partie de commande soit par suite de la différence de leurs gammes de fréquences, soit par suite de leur succession dans le temps. Pourtant, dans le cas général, cette distinction n'est pas de rigueur; une même action peut produire deux effets différents dont l'un prend en charge l'étude et l'autre, la conduite.

Dans les systèmes à commande duale les deux aspects décrits d'une action de commande donnent lieu à une contradiction. En effet, une commande ne peut être efficace que si son action est transmise au système en temps opportun. Mais une commande ne peut être efficace que lorsqu'on connaît avec une précision suffisante les propriétés du système. Or, l'étude d'un système demande du temps. Un organe de commande « hatif » produira des actions injustifiées, non étayées par l'information fournie par l'étude de l'objet; un système trop « prudent » temporisera en accumulant l'information et ne pourra pas guider l'objet en temps opportun vers le régime requis. Dans les deux cas le processus de commande peut ne pas être le meilleur et même inefficace.

Une information sur le système, bien qu'incomplète, est fournie par les répartitions probabilistes des caractéristiques. A mesure que se poursuit l'étude du système, ces répartitions définissent toujours mieux ses propriétés. C'est bien cette variation progressive des caractéristiques probabilistes a posteriori, qui s'approchent de plus

en plus des caractéristiques réelles qui permet d'estimer la vitesse d'étude du système. Le trait particulier des systèmes de commande duale est la dépendance de cette vitesse de la stratégie de l'organe de commande.

Enonçons maintenant le problème de construction d'un organe de commande optimal. A cette fin empruntons quelques concepts

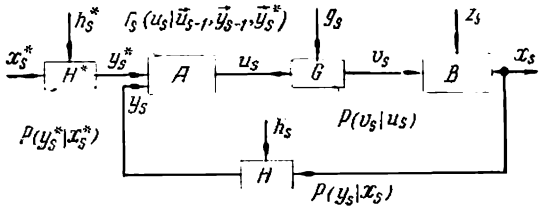


Fig. 6.2.

à la théorie des décisions statistiques. Formulons d'abord le problème de type particulier en le généralisant ensuite dans les paragraphes qui suivent.

Considérons le schéma représenté sur la figure 6.2 soumis aux conditions suivantes:

1) Le système étudié est un système continu à états discrets. Toutes les grandeurs qui figurent dans ce schéma sont envisagées à des instants discrets $t = 0, 1, \ldots, n$, où n est fixé. La valeur de ces grandeurs à l'instant $t = s$ est affectée de l'indice s (par exemple x_s^*, x_s, y_s, etc.).

2) On étudie le problème bayésien qui compte parmi ses données les densités de probabilité a priori. Soient h_s^*, h_s, g_s des suites des variables aléatoires indépendantes de densités de probabilité invariables $P(h_s^*)$, $P(h_s)$, $P(g_s)$ respectivement.

Ensuite, soit

$$z_s = z(s, \bar{\mu}), \tag{6.1}$$

où $\bar{\mu}$ est le vecteur aléatoire

$$\bar{\mu} = (\mu_1, \ldots, \mu_m). \tag{6.2}$$

Nous considérons que la densité de probabilité a priori $P(\bar{\mu})$ du vecteur $\bar{\mu}$ est imposée. Posons d'une manière analogue que la consigne x_s^* est de la forme

$$x_s^* = x_s^*(s, \bar{\lambda}), \tag{6.3}$$

où $\bar{\lambda}$ est le vecteur aléatoire

$$\bar{\lambda} = (\lambda_1, \ldots, \lambda_q), \tag{6.4}$$

de densité de probabilité a priori $P(\bar{\lambda})$ imposée. Admettons que toutes les actions extérieures z_s, x_s^*, h_s^*, h_s et g_s qui agissent sur le schéma sont statistiquement indépendantes.

3) Considérons que le système B est sans mémoire *). L'opérateur

$$x_s = F(z_s, v_s) \tag{6.5}$$

qui caractérise ce système associe les valeurs x_s, z_s, v_s se produisant au même instant. Posons que la fonction F est finie, univoque et dérivable mais pas forcément biunivoque.

4) Considérons que les modes de combinaison du signal et du bruit dans les blocs H^*, H et G sont connus et invariables et les blocs eux-mêmes sont sans mémoire. Par consequent,

$$v_s = v_s(u_s, g_s), \quad y_s^* = y_s^*(h_s^*, x_s^*), \quad y_s = y_s(h_s, x_s). \tag{6.6}$$

Au lieu d'imposer ces relations et les caractéristiques probabilistes des bruits h_s^*, h_s, g_s, on peut se donner directement les densités de probabilité conditionnelles $P(v_s \mid u_s)$, $P(y_s^* \mid x_s^*)$ et $P(y_s \mid x_s)$ (fig. 6.2).

5) Introduisons la fonction de pertes élémentaire

$$W_s = W_s(s, x_s, x_s^*). \tag{6.7}$$

Soit la fonction de pertes totale W qui s'écrit

$$W = \sum_{s=0}^{n} W_s(s, x_s, x_s^*). \tag{6.8}$$

Disons qu'un système est optimal s'il minimise le risque total R qui est l'espérance mathématique de la fonction de pertes

$$R = M\{W\} = \sum_{s=0}^{n} M\{W_s\} = \sum_{s=0}^{n} R_s. \tag{6.9}$$

Ici R_s est le risque élémentaire.

6) Considérons que dans le cas général l'organe de commande possède une mémoire. De plus, supposons pour la généralité que l'organe A réalise une stratégie aléatoire. Introduisons les vecteurs temporels

$$\left.\begin{array}{ll} \vec{u}_s = (u_0, u_1, \ldots, u_s), & \vec{x}_s^* = (x_0^*, x_1^*, \ldots, x_s^*), \\ \vec{v}_s = (v_0, v_1, \ldots, v_s), & \vec{y}_s^* = (y_0^*, y_1^*, \ldots, y_s^*), \\ \vec{x}_s = (x_0, x_1, \ldots, x_s), & \vec{y}_s = (y_0, y_1, \ldots, y_s) \\ & (0 \leqslant s \leqslant n). \end{array}\right\} \tag{6.10}$$

La valeur u_s est une fonction aléatoire des grandeurs y_i appliquées à l'entrée de l'organe A aux instants précédents $(i < s)$; u_s est également une fonction des y_j^* $(j \leqslant s)$; enfin, u_s peut dépendre aussi des valeurs u_ν enregistrées à la sortie de l'organe A dans le

*) Les systèmes de ce type sont d'un grand intérêt pour la pratique: si à chaque test le paramètre d'entrée v_s donne lieu à une valeur stationnaire x_s, le système dynamique se comporte comme un système sans mémoire.

passé ($\nu < s$). Les valeurs antérieures u_ν peuvent être mémorisées par cet organe et être appliquées à l'entrée d'un calculateur qui définit la valeur u_s courante.

Le problème consiste à définir la stratégie aléatoire optimale de l'organe A, c'est-à-dire les densités de probabilité optimales

$$P_s(u_s) = \Gamma_s(u_s \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s^*) \qquad (0 \leqslant s \leqslant n) \tag{6.11}$$

qui minimisent le risque total R. Etant donné que Γ_s est la densité de probabilité, $\Gamma_s \geqslant 0$ et la fonction Γ_s doivent vérifier la contrainte

$$\int\limits_{\Omega(u_s)} \Gamma_s(u_s)\, d\Omega = 1. \tag{6.12}$$

Appelons Γ_i ($i = 0, 1, \ldots, n$) stratégies élémentaires, et leur ensemble stratégie totale.

La première étape de résolution consiste à déduire la formule du risque R. Cherchons d'abord l'expression pour le risque conditionnel élémentaire r_s. Ce dernier peut être défini comme un terme de la formule du risque relatif au s-ième pas (c'est-à-dire à l'instant $t = s$), *les valeurs d'entrée de l'organe A étant fixées.* Ainsi, la formule de r_s se déduit en posant que les vecteurs temporels $\vec{y}_s^*$, $\vec{u}_{s-1}$, $\vec{y}_{s-1}$ sont donnés. Par conséquent,

$$r_s = M\{W_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}\} =$$
$$= \int\limits_{\Omega(\bar{\lambda}, x_s)} W_s[s, x_s^*(s, \bar{\lambda}), x_s] \cdot P(\bar{\lambda}, x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})\, d\Omega. \tag{6.13}$$

Explicitons cette formule. $\Omega(\bar{\lambda}, x_s)$ est ici le domaine qui comprend l'ensemble des valeurs possibles du vecteur $\bar{\lambda}$ et de la grandeur x_s. Ce domaine peut être représenté comme un espace de dimension $(q+1)$ aux coordonnées cartésiennes $\lambda_1, \lambda_2, \ldots, \lambda_q, x_s$. La notation $d\Omega$ de l'expression (6.13) est adoptée pour un élément infiniment petit du domaine $\Omega(\bar{\lambda}, x_s)$. Les grandeurs $\bar{\lambda}$ et x_s de l'expression de W_s étant aléatoires, la fonction W_s de la formule (6.13) de l'espérance mathématique doit être multipliée par la densité de probabilité conditionnelle conjointe de $\bar{\lambda}$ et de x_s, c'est-à-dire par la fonction $P(\bar{\lambda}, x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})$ sous la condition que $\vec{y}_s^*$, $\vec{u}_{s-1}$, $\vec{y}_{s-1}$ sont données. Le produit obtenu est intégré sur $\Omega(\bar{\lambda}, x_s)$; il en résulte l'expression du second membre de (6.13).

Cherchons l'expression de la densité de probabilité conditionnelle conjointe de $\bar{\lambda}$ et x_s. D'après le théorème du produit des probabilités [cf. (2.4)]

$$P(\bar{\lambda}, x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) =$$
$$= P(\bar{\lambda} \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) \cdot P(x_s \mid \bar{\lambda}, \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}). \tag{6.14}$$

Pourtant la densité de probabilité de $\overline{\lambda}$ n'est liée qu'à $\vec{y}_s^*$ (cf. fig. 6.2) et ne varie pas si, en plus du vecteur fixé $\vec{y}_s^*$, on fixe les vecteurs $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$. En effet, la chaîne qui comporte le bloc H^* de la figure 6.2 ne fait pas partie du schéma en boucle fermée et constitue un circuit à part. Les propriétés de cette chaîne permettent de définir la fonction $P(\overline{\lambda} \mid \vec{y}_s^*)$ qui ne dépend pas des valeurs $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$. Il vient donc

$$P(\overline{\lambda} \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) = P(\overline{\lambda} \mid \vec{y}_s^*) = P_s(\overline{\lambda}). \tag{6.15}$$

Nous avons employé ici la notation $P_s(\overline{\lambda})$ pour la fonction $P(\overline{\lambda} \mid \vec{y}_s^*)$. Cette expression est la densité de probabilité a posteriori de $\overline{\lambda}$ calculée à l'instant $t = s$ sur la base des valeurs observées y_0^*, $y_1^*, \ldots, y_s^*$.

On peut également simplifier le deuxième facteur du second membre de l'égalité (6.14). En effet, la densité de probabilité de x_s, les entrées $\vec{y}_s^*$, $\vec{y}_{s-1}$, $\vec{u}_{s-1}$ de l'organe A étant fixées, ne varie pas si on fixe encore $\overline{\lambda}$. Par conséquent,

$$P(x_s \mid \overline{\lambda}, \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) = P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}). \tag{6.16}$$

En portant (6.14) dans (6.13), compte tenu de (6.15) et (6.16), on obtient

$$r_s = \int\limits_{\Omega(\overline{\lambda},\, x_s)} W_s[s, x_s^*(s, \overline{\lambda}), x_s] \cdot P_s(\overline{\lambda}) \times \\ \times P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})\, d\Omega. \tag{6.17}$$

La figure 6.2 montre qu'ayant fixé $\overline{\mu}$ et donc z_s, ainsi que u_s, la sortie x_s du système est une grandeur aléatoire par suite de la présence d'une perturbation aléatoire g_s. Désignons par $P(x_s \mid \overline{\mu}, s, u_s)$ la densité de probabilité conditionnelle de x_s, avec $\overline{\mu}$ et u_s fixés, considérée à l'instant $t = s$. Soit également $P_s(\overline{\mu})$ la densité de probabilité a posteriori du vecteur $\overline{\mu}$ à l'instant $t = s$. Cette fonction diffère de la densité a priori $P(\overline{\mu}) = P_0(\overline{\mu})$ du vecteur $\overline{\mu}$, car l'information fournie à l'organe de commande A pendant le temps $t = 0, 1, \ldots, s-1$ permet de préciser la valeur du vecteur $\overline{\mu}$. La densité de probabilité $P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})$ est donnée par la formule

$$P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) = \\ = \int\limits_{\Omega(\overline{\mu},\, u_s)} P(x_s \mid \overline{\mu}, s, u_s) \cdot P_s(\overline{\mu})\, \Gamma_s(u_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})\, d\Omega. \tag{6.18}$$

Cette formule se vérifie aisément du fait que pour obtenir $P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})$ il faut prendre la moyenne de l'expression $P(x_s \mid \bar{\mu}, s, u_s)$ par rapport aux deux variables indépendantes $\bar{\mu}$ et u_s. Elles sont indépendantes pour peu que les entrées de l'organe de commande soient fixées. Cette condition est équivalente dans un certain sens à l'ouverture de la chaîne en boucle fermée du schéma représenté sur la figure 6.2 *).

En portant (6.18) dans (6.17), on a

$$r_s = \int\limits_{\Omega(\bar{\lambda}, \bar{\mu}, x_s, u_s)} W_s[s, x_s^*(s, \bar{\lambda}) x_s] \cdot P_s(\bar{\lambda}) \cdot P(x_s \mid \bar{\mu}, s, u_s) \times$$
$$\times P_s(\bar{\mu}) \cdot \Gamma_s(u_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})\, d\Omega. \quad (6.19)$$

Déterminons maintenant la densité de probabilité a posteriori $P_s(\bar{\lambda})$. Puisque

$$P(\bar{\lambda}, \vec{y}_s^*) = P(\bar{\lambda}) \cdot P(\vec{y}_s^* \mid \bar{\lambda}) = P(\bar{\lambda} \mid \vec{y}_s^*) \cdot P(\vec{y}_s^*), \quad (6.20)$$

il vient

$$P_s(\bar{\lambda}) = P(\bar{\lambda} \mid \vec{y}_s^*) = P(\bar{\lambda}) \frac{P(\vec{y}_s^* \mid \bar{\lambda})}{P(\vec{y}_s^*)}. \quad (6.21)$$

Ici $P(\bar{\lambda})$ est la densité de probabilité a priori de $\bar{\lambda}$, $P(\vec{y}_s^*)$ la densité de probabilité a priori (inconditionnelle) de $\vec{y}_s^*$ et $P(\vec{y}_s^* \mid \bar{\lambda})$ la densité de probabilité conditionnelle de $\vec{y}_s^*$, la valeur de $\bar{\lambda}$ étant fixée. Cette dernière fonction (dont la forme est celle d'une fonction de vraisemblance) peut s'obtenir si l'on connaît les propriétés du canal H^* (fig. 6.2).

Les valeurs h_i^* étant indépendantes pour les différents i et l'élément H^* étant sans inertie, on a **)

$$P(\vec{y}_s^* \mid \bar{\lambda}) = \prod_{i=0}^{s} P(y_i^* \mid \bar{\lambda}) = \prod_{i=0}^{s} P(y_i^* \mid i, \bar{\lambda}). \quad (6.22)$$

Il est beaucoup plus difficile de trouver la densité de probabilité a posteriori $P_s(\bar{\mu})$. Cette grandeur peut être calculée par l'organe A qui mémorise les valeurs u_i et y_i ($i = 0, \ldots, s-1$). On étudie ainsi l'entrée et la sortie du système B. Les grandeurs v_i d'entrée et x_i de sortie du système sont mesurées avec erreurs par suite de la présence des perturbations g_i et h_i. Toutefois, l'étude du système

*) La valeur x_s n'intervient pas directement dans $\vec{y}_{s-1}$ à travers le système, car ce dernier est sans mémoire. La liaison de x_s et $\vec{y}_{s-1}$ n'est assurée que par la stratégie de l'organe de commande A qui mémorise $\vec{y}_{s-1}$.

**) Nous avons souligné ici que P peut dépendre du numéro i du pas.

permet de préciser la densité de probabilité du vecteur des paramètres $\bar{\mu}$ qui définit z_s, perturbation subie par le système B. La perturbation z_s est en fait la caractéristique variable du système. La recherche de la meilleure méthode pour gouverner le système B est d'autant plus sûre que la connaissance du vecteur $\bar{\mu}$ et par conséquent de la caractéristique du système B est plus précise. C'est pourquoi il est très important de calculer la densité de probabilité a posteriori $P_s(\bar{\mu})$ qui recèle toute l'information sur le système emmagasinée par l'organe A.

Pour définir $P_s(\bar{\mu})$ servons-nous également de la formule du type bayésien. Considérons la densité de probabilité conjointe $P(\bar{\mu}, \vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*)$ des vecteurs $\bar{\mu}$, $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$ avec $\vec{y}_s^*$ fixé. Soit $P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*)$ la densité de probabilité conditionnelle conjointe des vecteurs $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$, $\bar{\mu}$ et, bien sûr, $\vec{y}_s^*$ étant fixés, et $P(\bar{\mu} \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s^*)$ la densité de probabilité conditionnelle de $\bar{\mu}$, les vecteurs $\vec{u}_{s-1}$, $\vec{y}_{s-1}$ et $\vec{y}_s^*$ étant fixés. Cette dernière fonction est précisément la densité de probabilité a posteriori recherchée du vecteur $\bar{\mu}$.

D'après le théorème du produit des probabilités

$$P(\bar{\mu}, \vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*) = P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*) \cdot P(\bar{\mu}) =$$
$$= P(\bar{\mu} \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s^*)\, P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*). \qquad (6.23)$$

Dans cette formule la fonction $P(\bar{\mu}) = P_0(\mu)$ est la densité de probabilité a priori (inconditionnelle) de $\bar{\mu}$ et $P(\vec{u}_{s-1}, \vec{y}_{s-1})$ la densité de probabilité inconditionnelle conjointe des vecteurs $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$. Comme $\bar{\mu}$ ne dépend pas de $\bar{\lambda}$ ni de $\vec{y}_s^*$, dans l'expression de $P(\bar{\mu})$ la dépendance de $\vec{y}_s^*$ est absente.

(6.23) entraîne

$$P_s(\bar{\mu}) = P(\bar{\mu} \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s^*) = \frac{P(\bar{\mu}) \cdot P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*)}{P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*)} =$$
$$= \frac{P(\bar{\mu}) \cdot P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*)}{\int\limits_{\Omega(\bar{\mu})} P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*)\, P(\bar{\mu})\, d\Omega}. \qquad (6.24)$$

Ainsi, pour définir $P_s(\bar{\mu})$ il faut trouver l'expression de la fonction de vraisemblance $P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*)$ dans un système en

boucle fermée représenté sur la figure 6.2. On voit aisément que la densité de probabilité d'un événement complexe qui consiste en l'apparition d'un couple de valeurs u_0, y_0 suivi de u_1, y_1, puis de u_2, y_2, etc., la valeur $\bar{\mu}$ étant fixée, est un produit de plusieurs facteurs : de la densité de probabilité $P(u_0, y_0 \mid \bar{\mu}, y_0^*)$ du premier de ces événements ; de la densité de probabilité $P(u_1, y_1 \mid \bar{\mu}, u_0, y_0, \vec{y}_1^*)$ du deuxième de ces événements sous la condition que le premier ait eu lieu ; de la densité de probabilité $P(u_2, y_2 \mid \bar{\mu}, \vec{u}_1, \vec{y}_1, \vec{y}_2^*)$ du troisième de ces événements sous la condition que le premier et le deuxième aient eu lieu, etc. :

$$P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*) = P(u_0, y_0 \mid \bar{\mu}, y_0^*) \times$$
$$\times P(u_1, y_1) \bar{\mu}, u_0, y_0, \vec{y}_1^*) P(u_2, y_2 \mid \bar{\mu}, \vec{u}_1, \vec{y}_1, \vec{y}_2^*) \ldots \times$$
$$\times \ldots P(u_{s-1}, y_{s-1} \mid \bar{\mu}, \vec{u}_{s-2}, \vec{y}_{s-2}, \vec{y}_{s-1}^*). \qquad (6.25)$$

Considérons l' i-ième facteur de cette expression $(0 < i \leqslant s-1)$ et réalisons sa transformation :

$$P(u_i, y_i \mid \bar{\mu}, \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{y}_i^*) =$$
$$= P(y_i \mid \bar{\mu}, u_i, \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{y}_i^*) P(u_i \mid \bar{\mu}, \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{y}_i^*). \qquad (6.26)$$

Le premier facteur du second membre de (6.26) est la densité de probabilité conditionnelle de y_i à l'instant $t = i$ sous la condition que $\bar{\mu}$ et u_i soient fixés. La figure 6.2 montre qu'elle ne change pas si l'on fixe également les vecteurs $\vec{y}_i^*$, $\vec{u}_{i-1}$ et $\vec{y}_{i-1}$; par conséquent, cette densité ne dépend pas des vecteurs $\vec{u}_{i-1}$, $\vec{y}_i^*$ et $\vec{y}_{i-1}$. On peut la désigner par $P(y_i \mid \bar{\mu}, i, u_i)$, i à droite du trait vertical signifiant que cette fonction dépend aussi du numéro i du pas. Ensuite, la densité de probabilité conditionnelle $P(u_i \mid \bar{\mu}, \vec{u}_{i-1}, \vec{y}_i, \vec{y}_i^*)$ ne dépend pas de $\bar{\mu}$ pour peu que $\vec{u}_{i-1}$ et $\vec{y}_{i-1}$ soient fixés. Cette fonction est une stratégie aléatoire Γ_i de l'organe de commande A. C'est pourquoi on peut l'écrire *) comme $\Gamma_i(u_i \mid \vec{y}_i^*, \vec{u}_{i-1}, \vec{y}_{i-1})$. Ainsi la formule (6.26) peut se mettre sous la forme

$$P(u_i, y_i \mid \bar{\mu}, \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{y}_i^*) = P(y_i \mid \bar{\mu}, i, u_i) \Gamma_i, \qquad (6.27)$$

où Γ_i est une notation abrégée de

$$\Gamma_i(u_i \mid \vec{y}_i^*, \vec{u}_{i-1}, \vec{y}_{i-1}).$$

*) Elle ne dépend pas de y_i si A assure la mémorisation des valeurs y_j $(j < i)$ antérieures. C'est pourquoi au lieu de $\vec{y}_i$ on peut écrire $\vec{y}_{i-1}$.

En portant (6.27) dans (6.26) et (6.26) dans (6.25) on obtient l'expression de la fonction de vraisemblance sous la forme

$$P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \bar{\mu}, \vec{y}_s^*) = [\prod_{i=0}^{s-1} P(y_i \mid \bar{\mu}, i, u_i)] \cdot \prod_{i=0}^{s-1} \Gamma_i, \tag{6.28}$$

dans laquelle on a introduit la notation

$$\Gamma_0 = P_0(u_0). \tag{6.29}$$

Cette densité de probabilité ne dépend pas des observations qui n'existaient pas encore à l'instant initial $t = 0$, elle n'est définie que par les données a priori.

La substitution de (6.28) dans l'expression (6.24) conduit à la formule

$$P_s(\bar{\mu}) = P(\bar{\mu} \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s^*) = \frac{P(\bar{\mu}) \cdot [\prod_{i=0}^{s-1} P(y_i \mid \bar{\mu}, i, u_i)] \cdot \prod_{i=0}^{s-1} \Gamma_i}{P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*)}. \tag{6.30}$$

Portons notre attention sur la différence de principe entre cette dernière formule et l'expression (6.21). L'accumulation de l'information sur les caractéristiques du système se traduit par le fait que la densité de probabilité a priori $P(\bar{\mu})$ est remplacée à chaque pas par des densités a posteriori $P_s(\bar{\mu})$ qui caractérisent le vecteur $\bar{\mu}$ avec une précision de plus en plus meilleure.

La formule (6.30) montre que la forme de la fonction $P_s(\bar{\mu})$ et donc la vitesse d'accumulation de l'information sur le système dépendent de toutes les stratégies antérieures Γ_i $(i < s)$. Autrement dit, la vitesse de l'étude du système dépend de l'organisation des expériences qui consistent à appliquer à ce système les actions u_i et à explorer les réponses y_i du système à ces actions.

Or, dans la formule (6.21) de $P_s(\bar{\lambda})$ ainsi que dans les formules analogues du chapitre précédent la dépendance entre la vitesse d'accumulation de l'information et la stratégie Γ_i $(i = 0, 1, \ldots, s-1)$ est absente. Cela veut dire que l'accumulation de l'information sur l'action x_s^* dans la chaîne en boucle ouverte est passive, ou indépendante.

La comparaison des formules (6.21) et (6.30) permet de rendre évidente la différence principale entre les commandes duale et non duale. Dans la commande duale la vitesse de l'étude du système est définie par le caractère des commandes, c'est-à-dire par la stratégie de l'organe de commande, alors que cette relation est inexistante dans le cas d'une commande non duale.

En portant les expressions (6.21) et (6.30) dans (6.19) on obtient la formule définitive du risque conditionnel élémentaire

$$r_s = \int\limits_{\Omega(\bar{\lambda},\ \bar{\mu},\ x_s,\ u_s)} W_s\,[s,\ x_s^*\,(s,\ \bar{\lambda}),\ x_s] \times$$

$$\times \frac{P\,(\bar{\lambda})\cdot P\,(\vec{y}_s^* \mid \bar{\lambda})}{P\,(\vec{y}_s^*)}\, P\,(x_s \mid \bar{\mu},\ s,\ u_s) \times \frac{P\,(\bar{\mu})\cdot \prod\limits_{i=0}^{s-1} P\,(y_i \mid \bar{\mu},\ i,\ u_i)}{P\,(\vec{y}_{s-1},\ \vec{u}_{s-1} \mid \vec{y}_s^*)} \cdot \prod_{i=0}^{s} \Gamma_i\, d\Omega. \tag{6.31}$$

Si l'on considère les valeurs de r_s fournies par différentes expériences, les vecteurs $\vec{y}_s^*$, $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$, en général inconnus d'avance, peuvent prendre différentes valeurs. Soit $P\,(\vec{y}_s^*,\ \vec{u}_{s-1},\ \vec{y}_{s-1})$ la densité de probabilité conjointe de ces vecteurs. Le risque élémentaire moyen R_s, qui représente la moyenne du risque conditionnel élémentaire r_s pour un très grand nombre d'expériences, est alors défini par la formule

$$R_s = \int\limits_{\Omega(\vec{y}_s^*,\ \vec{u}_{s-1},\ \vec{y}_{s-1})} r_s P\,(\vec{y}_s^*,\ \vec{u}_{s-1},\ \vec{y}_{s-1})\, d\Omega. \tag{6.32}$$

Tenons compte maintenant du fait que

$$P\,(\vec{y}_s^*,\ \vec{u}_{s-1},\ \vec{y}_{s-1}) = P\,(\vec{u}_{s-1},\ \vec{y}_{s-1} \mid \vec{y}_s^*)\, P\,(\vec{y}_s^*). \tag{6.33}$$

En portant (6.31) et (6.33) dans (6.22) et compte tenu de (6.22) on obtient la formule principale qui exprime R_s sous la forme suivante :

$$R_s = \int\limits_{\Omega(\bar{\lambda},\ \bar{\mu},\ x_s,\ \vec{y}_s^*,\ \vec{u}_s,\ \vec{y}_{s-1})} W_s\,[s,\ x_s^*\,(s,\ \bar{\lambda}),\ x_s] \times$$

$$\times P\,(\bar{\lambda}) \Big[\prod_{i=0}^{s} P\,(y_i^* \mid i,\ \bar{\lambda})\Big]\, P\,(x_s \mid \bar{\mu},\ s,\ u_s) \times$$

$$\times P\,(\bar{\mu}) \Big[\prod_{i=0}^{s-1} P\,(y_i \mid \bar{\mu},\ i,\ u_i)\Big] \Big[\prod_{i=0}^{s} \Gamma_i\Big]\, d\Omega. \tag{6.34}$$

Bien que dans le problème considéré le système B soit sans mémoire, le risque R_s associé à l'instant $t = s$ dépend de la stratégie totale, c'est-à-dire de l'ensemble des stratégies élémentaires qui sont les fonctions Γ_i aux instants $t = 0,\ 1,\ \ldots,\ s$. La cause physique de ce phénomène, inexistant dans un système en boucle

ouverte, est précisément la dualité de la commande. Il s'ensuit que la commande, pour $t = k$ $(k < n)$, doit avoir pour tâche non seulement la minimisation de la valeur R_k, risque élémentaire qui correspond à cet instant, mais aussi une meilleure étude du système pour contribuer à la diminution des valeurs des risques R_i $(i > k)$ aux instants ultérieurs. Considérons S_k, partie du risque total R dépendant de la stratégie Γ_k :

$$S_k = \sum_{i=k}^{n} R_i = R_k + \sum_{i=k+1}^{n} R_i. \tag{6.35}$$

Par rapport à la stratégie élémentaire Γ_k le premier terme du second membre de (6.35) peut s'appeler *risque d'action* et le deuxième *risque d'étude*. Une stratégie primitive qui choisit l'action u_k (ou sa densité de probabilité Γ_k) de façon à ne minimiser que le risque d'action n'est pas optimale. Par ailleurs, l'ignorance du risque d'action et la minimisation du seul risque d'étude, c'est-à-dire la sélection de u_k (ou de Γ_k) uniquement dans le but de mieux étudier le système pour utiliser l'information obtenue dans les actions ultérieures, ne correspond pas non plus à un comportement optimal. Dans le cas d'une commande duale, la stratégie est optimale si elle minimise la somme S_k des risques d'action et d'étude.

§ 2. Résolution du problème et exemples simples

Pour définir la stratégie optimale d'une commande duale aisons appel à la programmation dynamique [6.1, 6.7].

Introduisons les fonctions auxiliaires $(k = 0, 1, \ldots, n)$

$$\begin{aligned}\alpha_k = \alpha_k(\vec{u}_k, \vec{y}_{k-1}, \vec{y}_k^*) = \\ = \int\limits_{\Omega(\bar{\lambda}, \bar{u}, x_k)} W_k[k, x_k^*(k, \bar{\lambda}) x_k] \cdot P(\bar{\lambda}) \times \\ \times \Big[\prod_{i=0}^{k} P(y_i^* \mid i, \bar{\lambda})\Big] \cdot P(x_k \mid \bar{\mu}, k, u_k) \times \\ \times P(\bar{\mu}) \cdot \Big[\prod_{i=0}^{k-1} P(y_i \mid \bar{\mu}, i, u_i)\Big] d\Omega\end{aligned} \tag{6.36}$$

et

$$\beta_k = \prod_{i=0}^{k} \Gamma_k. \tag{6.37}$$

Quel est d'abord le risque R_n au dernier instant $t = n$, en considérant que les stratégies $\Gamma_0, \ldots, \Gamma_{n-1}$ sont fixées d'une façon ou d'une autre? Compte tenu de (6.37) et (6.36), (6.34) permet de tirer

$$R_n = \int\limits_{\Omega(\vec{u}_n,\ \vec{y}_{n-1},\ \vec{y}_n^*)} \alpha_n(\vec{u}_n, \vec{y}_{n-1}, \vec{y}_n^*) \cdot \beta_{n-1}\Gamma_n \, d\Omega =$$

$$= \int\limits_{\Omega(\vec{u}_{n-1},\ \vec{y}_{n-1},\ \vec{y}_n^*)} \beta_{n-1}\varkappa_n(\vec{u}_{n-1}, \vec{y}_{n-1}, \vec{y}_n^*)\, d\Omega, \quad (6.38)$$

où

$$\varkappa_n(\vec{u}_{n-1}, \vec{y}_{n-1}, \vec{y}_n^*) = \int\limits_{\Omega(\vec{u}_n)} \alpha_n(u_n, \vec{u}_{n-1}, \vec{y}_{n-1}, \vec{y}_n^*) \times$$

$$\times \Gamma_n(u_n \mid \vec{u}_{n-1}, \vec{y}_{n-1}, \vec{y}_n^*)\, d\Omega. \quad (6.39)$$

Le théorème de la moyenne et (6.12) nous autorisent à écrire

$$\varkappa_n = (\alpha_n)_{\text{moy}} \int\limits_{\Omega(\vec{u}_n)} \Gamma_n \, d\Omega = (\alpha_n)_{\text{moy}} \geqslant (\alpha_n)_{\min}. \quad (6.40)$$

Il faut choisir Γ_n telle qu'elle minimise R_n. Cela s'obtient si pour des $\vec{u}_{n-1}$, $\vec{y}_{n-1}$, $\vec{y}_n^*$ quelconques, Γ_n est choisie telle qu'elle minimise la fonction $\varkappa_n$. Dans ce but, de même que dans le chapitre précédent, cherchons $u_n = u_n^*$ minimisant la fonction α_n. Soit $\alpha_n = \gamma_n$ et

$$\gamma_n^* = \alpha_n(u_n^*, \vec{u}_{n-1}, \vec{y}_{n-1}, \vec{y}_n^*) =$$

$$= \min_{u_n \in \Omega(\vec{u}_n)} \alpha_n(u_n, \vec{u}_{n-1}, \vec{y}_{n-1}, \vec{y}_n^*). \quad (6.41)$$

La grandeur u_n^* est évidemment une fonction de $\vec{u}_{n-1}$, $\vec{y}_{n-1}$, $\vec{y}_n^*$:

$$u_n^* = u_n^*(\vec{u}_{n-1}, \vec{y}_{n-1}, \vec{y}_n^*). \quad (6.42)$$

La stratégie optimale Γ_n^* est donnée alors par l'expression

$$\Gamma_n^* = \delta(u_n - u_n^*), \quad (6.43)$$

où δ est l'impulsion unité. Cela signifie que Γ_n^* est une stratégie pure; de plus, d'après (6.42), la valeur optimale de u_n est égale à u_n^*. L'expression (6.42) montre que u_n^* est une fonction des valeurs u_s, y_s $(s = 0, 1, \ldots, n-1)$ fournies auparavant à l'organe de commande A, ainsi que des valeurs y_i^* $(i = 0, 1, \ldots, n)$.

On montre sans peine que la stratégie exprimée par les formules (6.41)-(6.43) est optimale. De même que dans le chapitre V,

portons l'expression (6.43) dans la relation (6.39). Il vient

$$\varkappa_n = (\alpha_n)_{u_n = u_n^*} = (\alpha_n)_{\min}. \tag{6.44}$$

Mais d'après (6.40) cette valeur est minimale possible pour $\varkappa_n$. Par conséquent Γ_n^* est une stratégie optimale.

Pour trouver les stratégies optimales Γ_i^* pour $i < n$ il faut rétrograder de l'instant terminal $t = n$ vers le départ. Pour déterminer Γ_{n-1}^*, par exemple, considérons les deux derniers termes qui font partie du risque total R, c'est-à-dire la grandeur

$$S_{n-1} = R_{n-1} + R_n = \int\limits_{\Omega(\vec{u}_{n-1},\, \vec{y}_{n-2},\, \vec{y}^*_{n-1})} \alpha_{n-1}\beta_{n-1}\, d\Omega +$$

$$+ \int\limits_{\Omega(\vec{u}_n,\, \vec{y}_{n-1},\, \vec{y}^*_n)} \alpha_n\beta_n\, d\Omega =$$

$$= \int\limits_{\Omega(\vec{u}_{n-2},\, \vec{y}_{n-2},\, \vec{y}^*_{n-1})} \beta_{n-2}\varkappa_{n-1}(\vec{u}_{n-2},\, \vec{y}_{n-2},\, \vec{y}^*_{n-1})\, d\Omega, \tag{6.45}$$

où

$$\varkappa_{n-1}(\vec{u}_{n-2},\, \vec{y}_{n-2},\, \vec{y}^*_{n-1}) = \int\limits_{\Omega(u_{n-1})} \Big\{\Gamma_{n-1}\alpha_{n-1} +$$

$$+ \Gamma_{n-1} \int\limits_{\Omega(y_{n-1},\, y^*_n)} \alpha^*_n\, d\Omega(y_{n-1},\, y^*_n)\Big\}\, d\Omega(u_{n-1}) =$$

$$= \int\limits_{\Omega(u_{n-1})} \Gamma_{n-1}\Big\{\alpha_{n-1} + \int\limits_{\Omega(y_{n-1},\, y^*_n)} \alpha^*_n\, d\Omega(y_{n-1},\, y^*_n)\Big\}\, d\Omega(u_{n-1}). \tag{6.46}$$

La fonction Γ_n n'est pas explicitée ici mais on sait de ce qui précède que pour ce faire il faut remplacer u_n par l'expression (6.42). Considérons qu'une telle substitution est déjà effectuée; Γ_n ne figure plus alors dans nos formules, tandis que u_n^* s'avère une fonction connue de $\vec{u}_{n-1}$, $\vec{y}_{n-1}$, $\vec{y}^*_n$. Examinons la fonction

$$\gamma_{n-1} = \gamma_{n-1}(\vec{u}_{n-1},\, \vec{y}_{n-2},\, \vec{y}^*_{n-1}) =$$

$$= \alpha_{n-1} + \int\limits_{\Omega(y_{n-1},\, y^*_n)} \gamma^*_n(u^*_n,\, \vec{u}_{n-1},\, \vec{y}_{n-1},\, \vec{y}^*_n)\, d\Omega. \tag{6.47}$$

Cherchons la valeur u^*_{n-1} qui minimise cette fonction. Il est clair que cette valeur dépend de toutes les autres variables qui

figurent dans l'expression de γ_{n-1} :

$$u^*_{n-1} = u^*_{n-1}(\vec{u}_{n-2}, \vec{y}_{n-2}, \vec{y}^*_{n-1}). \tag{6.48}$$

La stratégie optimale s'écrit alors

$$\Gamma^*_{n-1} = \delta(u_{n-1} - u^*_{n-1}), \tag{6.49}$$

ce qui signifie que Γ^*_{n-1} est également une stratégie pure. La validité de cette formule peut être prouvée tout comme cela a été fait plus haut pour Γ^*_n. En poursuivant cette procédure et en passant à la définition de $\Gamma_{n-2}, \ldots, \Gamma^*_0$ on aboutit à la règle suivante d'obtention de la stratégie optimale. Soit $\gamma_n = \alpha_n$,

$$\gamma_{n-k} = \alpha_{n-k} + \\ + \int\limits_{\Omega(y_{n-k},\, y^*_{n-k+1})} \gamma_{n-k+1}(u^*_{n-k+1}, \vec{u}_{n-k}, \vec{y}_{n-k}, \vec{y}^*_{n-k+1})\, d\Omega \tag{6.50}$$

et u^*_{n-k} est la valeur de u_{n-k} qui minimise la fonction γ_{n-k}, c'est-à-dire

$$\gamma^*_{n-k} = \min_{u_{n-k} \in \Omega(u_{n-k})} \gamma_{n-k} = \gamma_{n-k}(u^*_{n-k}). \tag{6.51}$$

Il est clair que

$$u^*_{n-k} = u^*_{n-k}(\vec{y}^*_{n-k}, \vec{u}_{n-k-1}, \vec{y}_{n-k-1}). \tag{6.52}$$

La stratégie optimale Γ^*_{n-k} est donnée par l'expression

$$\Gamma^*_{n-k} = \delta(u_{n-k} - u^*_{n-k}), \tag{6.53}$$

c'est-à-dire la stratégie optimale est une stratégie pure qui consiste à sélectionner $u_{n-k} = u^*_{n-k}$ d'après la formule (6.52). Comme on voit de cette formule, à l'instant $t = n - k$ la commande optimale u^*_{n-k} dépend de toute la « préhistoire » des grandeurs d'entrée appliquées à A, c'est-à-dire de toutes les valeurs u_i, y_i ($i = 0, 1, \ldots, n - k - 1$), ainsi que des valeurs y^*_j ($j \leqslant n - k$).

On montre facilement que dans l'expression de α_0 il faut poser

$$\prod_{i=0}^{n-1} P(y_i \mid \bar{\mu}, i, u_i) = 1. \tag{6.54}$$

On a alors

$$u^*_0 = u^*_0(y^*_0), \tag{6.55}$$

c'est-à-dire l'action de commande initiale dépend de la consigne fournie à l'entrée de l'organe de commande A et, de plus, bien entendu, des données a priori.

Il convient de noter que dans l'algorithme optimal (6.52) dont il faut munir l'organe de commande A, la valeur u^*_{n-k} dépend des grandeurs u_i, y_i, y^*_i appliquées à l'entrée de A dans le passé, ainsi que de la valeur courante y^*_{n-k}, mais non pas des valeurs futures de ces variables. C'est pourquoi un organe de commande optimal est physiquement réalisable.

Si les valeurs x^*_i $(i = 0, 1, \ldots, n)$ sont connues d'avance, la formule de la commande optimale se simplifie. C'est le cas des systèmes de stabilisation automatique ou des systèmes à programme. L'élément H^* de la figure 6.2 devient alors superflu et la loi de commande peut être munie d'avance de valeurs $x^*_i = y^*_i$ sous la forme de paramètres connus. Les formules pour les problèmes de ce type peuvent s'obtenir comme un cas particulier des formules déduites précédemment. Si l'on pose

$$P(\bar{\lambda}) = \delta(\bar{\lambda} - \bar{\lambda}^*), \quad P(y^*_i \mid i, \bar{\lambda}) = \delta\,[y^*_i - x^*_i(\bar{\lambda}^*)], \tag{6.56}$$

où $\bar{\lambda}$ est le vecteur fixé connu d'avance. Comme x^*_k peut également être considéré comme connu, la formule (6.34) prend la forme suivante

$$R_s = \int\limits_{\Omega(\bar{\mu},\, x_s,\, \vec{u}_s,\, \vec{y}_{s-1})} W_s(s, x^*_s, x_s) \cdot P(x_s \mid \bar{\mu}, s, u_s) \cdot P(\bar{\mu}) \times$$
$$\times \Big[\prod_{i=0}^{s-1} P(y_i \mid \bar{\mu}, i, u_i)\Big] \cdot \Big[\prod_{i=0}^{s} \Gamma_i\Big]\, d\Omega. \tag{6.57}$$

Si l'on pose maintenant, considérant x^*_k comme paramètre et ne l'explicitant pas, que

$$\alpha_k = \alpha_k(\vec{u}_k, \vec{y}_{k-1}) = \int\limits_{\Omega(\bar{\mu},\, x_k)} W_k(k, x^*_k, x_k) \cdot P(x_k \mid \bar{\mu}, k, u_k) \times$$
$$\times P(\bar{\mu}) \cdot \Big[\prod_{i=0}^{k-1} P(y_i \mid \bar{\mu}, i, u_i)\, d\Omega, \tag{6.58}$$

on aboutit à l'expression analogue à (6.38) :

$$R_k = \int\limits_{\Omega(\vec{u}_k,\, \vec{y}_{k-1})} \alpha_k(\vec{u}_k, \vec{y}_{k-1})\, \beta_k\, d\Omega. \tag{6.59}$$

En appliquant à ce cas les mêmes raisonnements que ceux utilisés dans ce qui précède, on tombe sur la procédure suivante de la recherche d'une stratégie optimale ; soit $\gamma_n = \alpha_n$ et

$$\gamma_{n-k} = \alpha_{n-k} + \int\limits_{\Omega(y_{n-k})} \gamma_{n-k+1}(u^*_{n-k+1}, u_{n-k}, \vec{y}_{n-k})\, d\Omega. \tag{6.60}$$

Soit, ensuite, u^*_{n-k} la valeur de u_{n-k} qui minimise la fonction γ_{n-k} d'après l'expression (6.51). Il est évident que

$$u^*_{n-k} = u^*_{n-k}(\vec{u}_{n-k-1}, \vec{y}_{n-k-1}). \tag{6.61}$$

Le paramètre $\vec{x}^*_n$ n'est pas présent ici explicitement. La stratégie optimale Γ^*_{n-k} est donnée alors par l'expression (6.53) où, toutefois, u^*_{n-k} se calcule d'après les relations (6.60) et (6.61).

Examinons d'abord certains exemples relativement simples donnés par [6.1, 6.7]. La figure 6.3 représente le schéma fonctionnel d'un système très simple. La grandeur de sortie x_s de l'objet commandé B est appliquée à une entrée de l'organe de commande A.

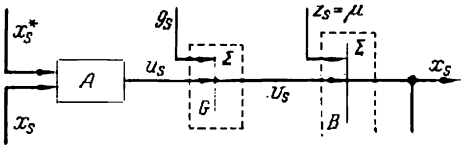


Fig. 6.3.

La consigne x^*_s, supposée connue, est fournie à une autre entrée de cet organe. La grandeur de sortie u_s de l'organe A est une commande appliquée à l'entrée de l'objet B par la chaîne d'action G où le signal de commande se mélange au bruit g_s. Dans cet exemple la perturbation z_s se ramène à une grandeur aléatoire μ. Le système B a un opérateur du type

$$x_s = v_s + \mu, \tag{6.62}$$

où

$$v_s = u_s + g_s. \tag{6.63}$$

Cette dernière expression est une loi de composition additive du bruit et du signal dans la chaîne d'action G. La fonction de pertes élémentaire est définie par l'expression

$$W_s = (x_s - x^*_s)^2. \tag{6.64}$$

Soient $P_0(\mu)$ et $P(g_s) = q(g_s)$ les distributions normales de moyennes nulles et de variances σ^2_μ et σ^2_g respectivement:

$$\left.\begin{aligned} P_0(\mu) &= \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{\frac{\mu^2}{2\sigma^2_\mu}\right\}, \\ P(g_s) = q(g_s) &= \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{g^2_s}{2\sigma^2_g}\right\}. \end{aligned}\right\} \tag{6.65}$$

Les valeurs de ces variables ne sont restreintes par aucune contrainte.

Il faut réaliser la synthèse d'un organe de commande optimal A minimisant le risque R, espérance mathématique de la fonction de pertes $W = \sum_{s=0}^{n} W_s$.

Selon la théorie générale il faut trouver d'abord la fonction α_k définie par (6.58). A cette fin on doit connaître au préalable les expressions $P(x_k \mid \bar{\mu}, k, u_k)$ et $P(y_i \mid \bar{\mu}, i, u_i)$. Puisque dans le problème concerné $y = x$, μ est une grandeur scalaire et la densité de probabilité ne dépend pas de i ou de k, on la désigne par $P(x_i \mid \mu, u_i)$.

(6.62) et (6.63) entraînent

$$x_i = (u_i + \mu) + g_i. \tag{6.66}$$

Par conséquent,

$$P(x_i \mid \mu, u_i) = q(x_i - u_i - \mu), \tag{6.67}$$

q étant la densité du bruit g_s donnée par la dernière des expressions (6.65). En portant cette expression dans (6.58), on a

$$\alpha_k = \int\limits_{\Omega(x_k, \mu)} W_k P_0(\mu) \cdot \prod_{i=0}^{k} P(x_i \mid \mu, u_i)\, d\Omega =$$

$$= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (x_k^* - x_k)^2 \cdot P_0(\mu) \prod_{i=0}^{k} q(x_i - u_i - \mu)\, dx_k\, d\mu. \tag{6.68}$$

Si l'on porte dans cette relation (6.65) on obtient

$$\alpha_k = \frac{1}{(\sigma_g)^{k+1} \sigma_\mu (2\pi)^{\frac{k}{2}+1}} \int\limits_{\mu=-\infty}^{\infty} \exp\left\{ -\frac{\mu^2}{2\sigma_\mu^2} - \sum_{i=0}^{k-1} \frac{(x_i - u_i - \mu)^2}{2\sigma_g^2} \right\} \times$$

$$\times \left[\int\limits_{x_k=-\infty}^{\infty} (x_k^* - x_k)^2 \exp\left\{ -\frac{(x_k - u_k - \mu)^2}{2\sigma_g^2} \right\} dx_k \right] d\mu. \tag{6.69}$$

Utilisons la formule ($p > 0$)

$$\int\limits_{-\infty}^{\infty} x^2 \exp\{-px^2 + 2qx\}\, dx = \frac{1}{2p} \sqrt{\frac{\pi}{p}} \left(1 + \frac{2q^2}{p}\right) \exp\left\{\frac{q^2}{p}\right\}. \tag{6.70}$$

Calculons l'intégrale I_k entre crochets de (6.69)

$$I_k = \exp\left\{ -\frac{l_k^2}{2\sigma_g^2} \right\} \int\limits_{-\infty}^{\infty} z^2 \exp\left\{ -\frac{z^2}{2\sigma_g^2} + 2\left(-\frac{l_k}{2\sigma_g^2}\right)^2 z \right\} dz, \tag{6.71}$$

avec

$$x_k - x_k^* = z, \quad x_k - u_k - \mu = l_k. \tag{6.72}$$

(6.70) permet d'obtenir à partir de (6.71) et (6.72)

$$I_k = \sigma_g \sqrt{2\pi}\,[\sigma_g^2 + (x_k^* - u_k - \mu)^2]. \tag{6.73}$$

On a alors en vertu de (6.69)

$$\alpha_k = \frac{1}{\sigma_g^k \cdot \sigma_\mu (2\pi)^{\frac{k+1}{2}}} \Big[\sigma_g^2 \int_{-\infty}^{\infty} \exp\Big\{-\frac{\mu^2}{2\sigma_\mu^2} - \sum_{i=0}^{k-1} \frac{(x_i - u_i)^2 - 2\mu(x_i - u_i) + \mu^2}{2\sigma_g^2}\Big\} d\mu + \int_{-\infty}^{\infty} \exp\Big\{-\frac{\mu^2}{2\sigma_\mu^2} - \sum_{i=0}^{k-1} \frac{(x_i - u_i)^2 - 2\mu(x_i - u_i) + \mu^2}{2\sigma_g^2}\Big\} (x^* - u_k - \mu)^2\, d\mu\Big]. \tag{6.74}$$

Pour calculer la première des intégrales entre crochets de l'expression (6.74), on peut faire appel à la formule ($p > 0$)

$$\int_{-\infty}^{\infty} \exp\{-px^2 \pm qx\}\, dx = \sqrt{\frac{\pi}{p}} \exp\Big\{\frac{q^2}{4p}\Big\}. \tag{6.75}$$

La deuxième intégrale entre crochets se définit en posant $\mu + u_k - x^* = z$ et en appliquant la formule (6.70). Il en résulte la forme définitive de la formule de α_k. Pour l'obtenir introduisons les notations

$$\left.\begin{aligned} & u_k - x_k^* = w_k, \quad \sum_{i=0}^{k-1} (x_i - u_i) = \Sigma_{k-1}, \\ & \varepsilon_k = \frac{1}{2\sigma_\mu^2} + \frac{k}{2\sigma_g^2}, \quad \sum_{i=0}^{k-1} (x_i - u_i)^2 = \theta_{k-1}, \\ & a_k = \frac{\sigma_g^2}{\sqrt{2}\,\sigma_g^k \cdot \sigma_\mu (2\pi)^{k/2} \sqrt{\varepsilon_k}}, \quad b_k = \frac{1}{2\sigma_g^k \sigma_\mu (2\pi)^{k/2} (2\varepsilon_k)^{3/2}}, \end{aligned}\right\} \tag{6.76}$$

ce qui donne

$$\alpha_k = \Big\{a_k + b_k + \frac{2b_k}{\varepsilon_k}\Big(w_k \varepsilon_k + \frac{\Sigma_{k-1}}{2\sigma_g^2}\Big)^2\Big\} \times \exp\Big\{-\frac{\theta_{k-1}}{2\sigma_g^2} + \frac{\Sigma_{k-1}^2}{4\sigma_g^2 \varepsilon_k}\Big\}. \tag{6.77}$$

Dans cette formule u_k, qui fait partie de l'expression de w_k, n'est présente que dans le facteur placé devant l'exponentielle, mais ni cette variable ni sa fonction ne figurent dans l'exposant. Cette circonstance rend beaucoup plus simple la minimisation de α_k par rapport à u_k.

Calculons maintenant Γ_i^* en procédant pas à pas. En posant dans la formule (6.77) $k = n$, minimisons la fonction $\gamma_n = \alpha_n$ par rapport à u_n. La condition du minimum coïncide avec l'annulation de l'expression entre parenthèses de (6.77):

$$w_n \varepsilon_n + \frac{\Sigma_{n-1}}{2\sigma_g^2} = 0. \tag{6.78}$$

En remplaçant les valeurs de w_n et Σ_{n-1} d'après (6.76) on en tire la commande optimale:

$$u_n^* = x_n^* - \frac{\sum\limits_{i=0}^{n-1} (x_i - u_i)}{2\sigma_g^2 \varepsilon_n}. \tag{6.79}$$

Ensuite, d'après (6.77) on définit

$$\gamma_n^* = \alpha_n^* = (a_n + b_n) \exp\left\{ -\frac{\theta_{n-1}}{2\sigma_g^2} + \frac{\Sigma_{n-1}^2}{4\sigma_g^4 \varepsilon_n} \right\}. \tag{6.80}$$

On calcule enfin la fonction γ_{n-1}:

$$\gamma_{n-1} = \alpha_{n-1} + \int\limits_{-\infty}^{\infty} \alpha_n^* \, dx_{n-1} = \left\{ \alpha_{n-1} + b_{n-1} + \right.$$

$$\left. + \frac{2b_{n-1}}{\varepsilon_{n-1}} \left[w_{n-1}\varepsilon_{n-1} + \frac{\Sigma_{n-2}}{2\sigma_g^2} \right]^2 \right\} \exp\left\{ -\frac{\theta_{n-2}}{2\sigma_g^2} + \frac{\Sigma_{n-2}^2}{4\sigma_g^4 \varepsilon_{n-1}} \right\} +$$

$$+ \int\limits_{-\infty}^{\infty} (a_n + b_n) \exp\left\{ -\frac{(x_{n-1} - u_{n-1})^2 + \theta_{n-2}}{2\sigma_g^2} + \right.$$

$$\left. + \frac{\Sigma_{n-2} + (x_{n-1} - u_{n-1})^2}{4\sigma_g^4 \varepsilon_n} \right\} dx_{n-1}. \tag{6.81}$$

Dans cette formule l'intégrale ne dépend pas de u_{n-1}. En effet, u_{n-1} n'est présent dans α_n^* que sous la forme de la différence $(x_{n-1} - u_{n-1})$. C'est pourquoi en remplaçant x_{n-1} par une nouvelle variable $x_{n-1} - u_{n-1} = z$, $dx_{n-1} = dz$, on obtient une intégrale dans laquelle ni l'expression sous le signe somme ni les limites d'intégration $(-\infty, \infty)$ ne dépendent de u_{n-1}.

Par conséquent, dans (6.81) seul α_{n-1} dépend de u_{n-1}. Cela signifie que pour $t = n - 1$ le risque d'action dépend, comme il fallait s'y attendre d'ailleurs, de la commande u_{n-1}, alors que le risque d'étude est indépendant de u_{n-1}, bien qu'il diffère du zéro. Les systèmes dans lesquels le processus de l'étude du système est le même quelle que soit la commande sont dits neutres [6.1]. Le système considéré est justement un système neutre. Le sens physique de ce phénomène dans l'exemple considéré est expliqué dans ce qui suit.

La minimisation de α_{n-1} par rapport à u_{n-1} donne la condition du minimum sous la forme de l'égalité à zéro de l'expression entre parenthèses de (6.77) :

$$w_{n-1}\varepsilon_{n-1}+\frac{\Sigma_{n-2}}{2\sigma_g^2}=0\,; \tag{6.82}$$

on en tire

$$u_{n-1}^*=x_{n-1}^*-\frac{\sum\limits_{i=0}^{n-2}(x_i-u_i)}{2\sigma_g^2\varepsilon_{n-1}}. \tag{6.83}$$

Des raisonnements complètement analogues conduisent à la conclusion que γ_s^* ne dépend pas non plus de w_s mais dépend seulement de la différence $x_{s-1}-u_{s-1}$. Aussi, en remplaçant $z=x_{s-1}-u_{s-1}$, on élimine dans l'intégrale $\int \gamma_s^*\,dx_{s-1}$ la dépendance de u_{s-1}. Par conséquent, quel que soit $(s-1)$ dans la formule de γ_{s-1}, il n'y a que le premier terme α_{s-1} qui dépend de u_{s-1}. A chaque pas le système est neutre. On en déduit que la formule de u_s coïncide avec (6.79), et (6.83) pour un s quelconque donne :

$$u_s^*=x_s^*-\frac{\sum\limits_{i=0}^{s-1}(x_i-u_i)}{2\sigma_g^2\varepsilon_s}\qquad (s=0,\ 1,\ \ldots,\ n). \tag{6.84}$$

En remplaçant ε_s par sa valeur tirée de (6.76) on aboutit à la loi de commande optimale suivante :

$$u_s^*=x_s^*-\frac{\sum\limits_{i=0}^{s-1}(x_i-u_i)}{s+\left(\frac{\sigma_g}{\sigma_\mu}\right)^2}\qquad (s=1,\ 2,\ \ldots,\ n). \tag{6.85}$$

Pour $s=0$ on obtient $u_0^*=x_0^*$.

Le sens de la formule (6.85) est bien simple. Si la valeur de μ était connue, il faudrait établir la grandeur $u_s=x_s^*-\mu$ pour satisfaire à la condition idéale $x_s=x_s^*$ (en l'absence du bruit g_s). Le deuxième terme de (6.85) donne, par conséquent, l'estimation de μ. S'il n'y avait pas de bruit ($g_s=0$), il serait facile de définir pour $u_s=v_s$ la valeur de μ : $\mu=x_i-u_i$. Pourtant, la figure 6.4 montre que

$$x_i-u_i=(\mu+v_i)-u_i=\mu+g_i. \tag{6.86}$$

Par conséquent, la différence x_i-u_i donne la valeur de μ mesurée avec une erreur g_i. Ainsi, l'estimation de μ s'obtient à partir de plusieurs mesures des différences x_i-u_i pour $i=0,1,\ldots,s-1$. Pour des s assez grands, le deuxième terme (6.85) devient la moyenne

arithmétique des $x_i - u_i$ pour $i = 0, 1, \ldots, s - 1$, ce qui correspond à des idées intuitives.

Attirons l'attention sur la ressemblance qui existe entre la formule (6.85) et l'expression (5.157) du chapitre précédent. Si l'on pose $\mu_0 = 0$ (ce qui a été supposé dans le dernier problème), ces expressions deviennent identiques. Il en est ainsi parce que le schéma 6.3 est réductible. Il peut être ramené à celui de la figure 5.17.

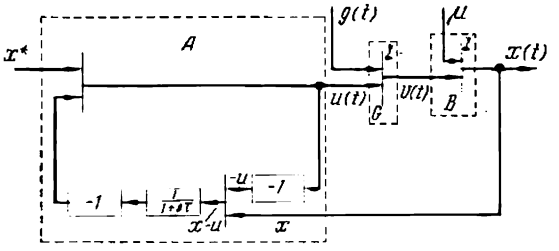


Fig. 6.4.

En effet, la formule (6.86) entraîne que mesurer x_s et u_s est équivalent à mesurer la perturbation $z_s = \mu$ avec une erreur g_s. Si $e_s = g_s$, c'est-à-dire si $\sigma_e = \sigma_g$, d'un certain point de vue, les schémas 5.17 et 6.3 sont équivalents. L'information fournie à l'organe de commande est la même dans les deux cas.

L'exemple ci-dessus montre que dans un système en boucle fermée également la vitesse d'accumulation de l'information sur le système peut être indépendante des valeurs de la commande. Les systèmes de ce type s'appellent *neutres*. Ainsi, le système représenté sur la figure 6.3 est neutre. Bien entendu, tous les systèmes réductibles sont neutres du fait que, dans les systèmes en boucle ouverte qui leur sont équivalents, l'accumulation de l'information sur le système ne dépend pas de la stratégie de l'organe de commande.

L'extrapolation des résultats de l'exemple ci-dessus au cas d'un système continu présente un certain intérêt. Soit la durée du pas

$$\Delta t = \frac{t}{s};$$

faisons tendre s vers l'infini en maintenant $t = \text{const}$; Δt tend alors vers zéro. Supposons que le problème consiste à construire un système de commande pour le cas où $g(t)$ est un bruit blanc stationnaire de densité spectrale S_0. On peut montrer que si l'on pose

$$\sigma_g^2 = \frac{S_0}{\Delta t}, \tag{6.87}$$

la suite des variables normales indépendantes g_s devient à la limite un bruit de ce type.

Récrivons la formule (6.85) compte tenu de (6.87)

$$u_s^* = x_s^* - \frac{\sum_{i=0}^{s-1} (x_i - u_i)\,\Delta t}{s\Delta t + \frac{\sigma_g^2 \Delta t}{\sigma_\mu^2}}. \tag{6.88}$$

Etant donné que d'après (6.87) $\sigma_g^2 \Delta t = S_0$ et que $s\Delta t = t$, on peut obtenir à la limite à partir de (6.88) pour $\Delta t \to 0$

$$u^*(t) = x^*(t) - \frac{\int_0^t (x-u)\,dt}{t + \frac{S_0}{\sigma_\mu^2}} = x^*(t) - \frac{\int_0^t (x-u)\,dt}{t+a}, \tag{6.89}$$

où $a = \frac{S_0}{\sigma_\mu^2}$. Le deuxième terme de cette expression peut s'obtenir à la sortie d'un filtre de fonction de transfert ($u_{\text{ent}} = 1$ pour $t > 0$)

$$u_{\text{sor}} = \frac{1}{t+a} \int_0^t u_{\text{ent}}\,dt = \frac{t}{t+a}, \tag{6.90}$$

si on applique à l'entrée du filtre la différence $x - u$. La fonction de transfert (6.90) n'est pas une exponentielle ; toutefois, on peut l

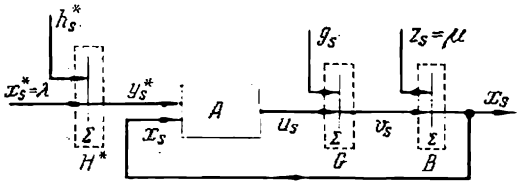


Fig. 6.5.

remplacer approximativement par une exponentielle à constante de temps $\tau \approx 1{,}1\ a$. Le système voisin de l'optimal est alors celui de la figure 6.4.

Considérons un autre exemple, un peu plus complexe, lorsqu'on accumule simultanément l'information sur la consinge x_s^* et la perturbation z_s subie par l'objet B. La figure 6.5 donne un schéma dont les blocs B et G sont les mêmes que sur la figure 6.3, mais la consigne est canalisée vers l'organe de commande A par la chaîne H^* où elle est entachée d'un bruit h_s^*. Les équations des chaînes H^*

et G s'écrivent respectivement

$$y_s^* = x_s^* + h_s^* \tag{6.91}$$

et

$$v_s = u_s + g_s. \tag{6.92}$$

L'équation du système B est de la forme

$$x_s = v_s + \mu = u_s + g_s + \mu. \tag{6.93}$$

Soit

$$x_s^* = \lambda = \text{const}, \quad z_s = \mu = \text{const}, \tag{6.94}$$

et, de plus, les variables λ, μ, g_s et h_s^* sont normales :

$$\left.\begin{aligned} P(\mu) &= \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{\mu^2}{2\sigma_\mu^2}\right\}, \\ P(\lambda) &= \frac{1}{\sigma_\lambda \sqrt{2\pi}} \exp\left\{-\frac{(\lambda-\lambda_0)^2}{2\sigma_\lambda^2}\right\}, \\ P(g_s) &= q(g_s) = \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{g_s^2}{2\sigma_g^2}\right\}, \\ P(h_s^*) &= \rho(h_s^*) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{h_s^{*2}}{2\sigma_h^2}\right\}. \end{aligned}\right| \tag{6.95}$$

La fonction de pertes élémentaire est donnée par l'expression

$$W_s = (x_s - x_s^*)^2 = (x_s - \lambda)^2. \tag{6.96}$$

Cherchons la stratégie optimale de l'organe de commande A et à cet effet calculons au préalable, d'après (6.36), la fonction

$$\alpha_k = \int\limits_{\Omega(\lambda,\, \mu,\, x_k)} (x_k - \lambda)^2 P(\lambda) \left[\prod_{i=0}^{k} P(y_i^* \mid \lambda)\right] \cdot P(\mu) \times$$
$$\times \left[\prod_{i=0}^{k} P(x_i \mid \mu, u_i)\right] d\Omega. \tag{6.97}$$

Dans le problème considéré $y_i = x_i$ et, de même que dans la formule (6.67) de l'exemple précédent,

$$P(x_i \mid \mu, u_i) = q(x_i - u_i - \mu) = \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{(x_i - u_i - \mu)^2}{2\sigma_g}\right\}. \tag{6.98}$$

En vertu de (6.91) on a

$$h_s^* = y_s^* - x_s^* = y_s^* - \lambda. \tag{6.99}$$

Donc

$$P(y_i^* \mid \lambda) = \rho(y_i - \lambda) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{(y_i^* - \lambda)^2}{2\sigma_h^2}\right\}. \quad (6.100)$$

En portant (6.98) et (6.100) dans (6.97), il vient

$$\alpha_k = \int\limits_{\Omega(\lambda,\,\mu)} P(\lambda) \cdot P(\mu) \left[\prod_{i=0}^{k} \rho(y_i^* - \lambda)\right] \cdot \left[\prod_{i=0}^{k-1} q(x_i - u_i - \mu)\right] \times$$
$$\times \left[\int\limits_{\Omega(x_k)} (x_k - \lambda)^2 q(x_k - u_k - \mu)\, d\Omega(x_k)\right] d\Omega(\lambda, \mu). \quad (6.101)$$

Calculons l'intégrale entre crochets de l'expression (6.101) :

$$I_k = \int\limits_{-\infty}^{\infty} (x_k - \lambda)^2 q(x_k - u_k - \mu)\, dx_k =$$
$$= \sigma_g \sqrt{2\pi}\, [\sigma_g^2 + (\lambda - u_k - \mu)^2]. \quad (6.102)$$

Cette intégrale se calcule de même que l'intégrale (6.73). Par conséquent, d'après (6.101) et (6.102)

$$\alpha_k = \int\limits_{-\infty}^{\infty} P(\lambda) \left[\prod_{i=0}^{k} \rho(y_i^* - \lambda)\right] \left\{a_k + b_k + \frac{2b_k}{\varepsilon_k} \times\right.$$
$$\left.\times \left(w_k \varepsilon_k + \frac{\Sigma_{k-1}}{2\sigma_g^2}\right)^2\right\} \exp\left\{-\frac{\theta_{k-1}}{2\sigma_g^2} + \frac{\Sigma_{k-1}^2}{4\sigma_g^4 \varepsilon_k}\right\} d\lambda. \quad (6.103)$$

Les notations employées ici sont les mêmes que celles des formules (6.76) et (6.77).

Minimisons la formule pour α_n par rapport à u_n qui n'est contenu que dans le terme w_n. Il se trouve que u_n^* de même que $\gamma_n^* = \alpha_n^* = \min \alpha_n$ dépendent des grandeurs $u_0, u_1, \ldots, u_{n-1}$; de plus, u_{n-1} est contenu dans les termes Σ_{n-1} et θ_{n-1}. Toutefois, dans ces termes u_{n-1} dépend seulement de la différence $x_{n-1} - u_{n-1} = \nu_{n-1}$. Par conséquent, u_n^* et γ_n^* dépendent précisément de cette différence. En intégrant ensuite par rapport à x_{n-1}, on peut remplacer cette variable par une autre variable ν_{n-1}. Alors, dans l'intégrale

$$\int\limits_{-\infty}^{\infty} \gamma_n^*(x_{n-1} - u_{n-1})\, dx_{n-1} = \int\limits_{-\infty}^{\infty} \gamma_n^*(\nu_{n-1})\, d\nu_{n-1} \quad (6.104)$$

la dépendance de u_{n-1} disparaît. Par conséquent, le deuxième terme de l'expression (6.60) pour γ_{n-1} ne dépend pas de u_{n-1} ; dans ce cas il n'y a que α_{n-1} qui dépend de u_{n-1}. Au lieu de minimiser γ_{n-1} par rapport à u_{n-1} il suffit donc de minimiser α_{n-1} par rapport

à u_{n-1} et obtenir u^*_{n-1}. En poursuivant ce raisonnement on peut arriver à la conclusion que tout u^*_s peut être trouvé en minimisant α_s par rapport à u_s. D'après la formule (6.103) pour α_s, il suffit d'étudier sa partie qui dépend de $w_s = u_s - \lambda$, et notamment :

$$J_s = \int_{-\infty}^{\infty} \left(u_s\varepsilon_s + \frac{\Sigma_{s-1}}{2\sigma_g^2} - \varepsilon_s\lambda\right)^2 P(\lambda) \left[\prod_{i=0}^{s} \rho(y_i^* - \lambda)\right] d\lambda. \quad (6.105)$$

Posons

$$J_{m,s} = \int_{-\infty}^{\infty} \lambda^m P(\lambda) \left[\prod_{i=0}^{s} \rho(y_i^* - \lambda)\right] d\lambda. \quad (6.106)$$

Il vient

$$J_s = \left(u_s\varepsilon_s + \frac{\Sigma_{s-1}}{2\sigma_g^2}\right) J_{0,s} - 2\varepsilon_s\left(u_s\varepsilon_s + \frac{\Sigma_{s-1}}{2\sigma_g^2}\right) J_{1,s} + \varepsilon_s^2 J_{2,s}. \quad (6.107)$$

La condition

$$\frac{dJ_s}{dv_s} = 0 \quad (6.108)$$

donne la valeur u_s^* minimisant J_s :

$$u_s^* = \frac{J_{1,s}}{J_{0,s}} - \frac{\Sigma_{s-1}}{2\varepsilon_s\sigma_g^2} = \frac{J_{1,s}}{J_{0,s}} - \frac{\sum_{i=0}^{s-1}(x_i - u_i)}{s + \left(\frac{\sigma_g}{\sigma_\mu}\right)^2}. \quad (6.109)$$

En divisant le numérateur et le dénominateur du premier terme de la formule (6.109) par $P(\vec{y}_s^*)$ [comparer avec (6.21)], on voit aisément que ce terme est l'espérance mathématique de la grandeur λ à l'instant $t = s$

$$\frac{J_{1,s}}{J_{0,s}} = \frac{\int_{-\infty}^{\infty} \lambda P_s(\lambda)\, d\lambda}{\int_{-\infty}^{\infty} P_s(\lambda)\, d\lambda} = \int_{-\infty}^{\infty} \lambda P_s(\lambda)\, d\lambda = M\{\lambda \mid \vec{y}_s^*\}. \quad (6.110)$$

L'intégration à l'aide des formules (6.70) et (6.75) conduit à

$$M\{\lambda \mid \vec{y}_s^*\} = \frac{J_{1,s}}{J_{0,s}} = \frac{\lambda_0}{1 + \left(\frac{\sigma_\lambda}{\sigma_h}\right)^2 (s+1)} + \frac{\sum_{r=0}^{s} y_r^*}{\left(\frac{\sigma_h}{\sigma_\lambda}\right)^2 + s + 1}. \quad (6.111)$$

En portant (6.111) dans (6.109) on aboutit à la formule définitive qui détermine la stratégie optimale

$$u_s^* = \frac{\lambda_0}{1+\left(\frac{\sigma_\lambda}{\sigma_h}\right)^2(s+1)} + \frac{\sum\limits_{r=0}^{s} y_r^*}{\left(\frac{\sigma_h}{\sigma_\lambda}\right)^2+(s+1)} - \frac{\sum\limits_{i=0}^{s-1}(x_i-u_i)}{s+\left(\frac{\sigma_g}{\sigma_\mu}\right)^2}. \quad (6.112)$$

Si les bruits g_s et h_s^* n'existaient pas, pour assurer la condition idéale $x_s = x_s^*$ il faudrait choisir $u_s = x_s^* - \mu = \lambda - \mu$. Les deux premiers termes de la formule (6.112) donnent l'estimation de λ, et le dernier celle de μ. Pour les petits s, la moyenne a priori de λ_0 peut jouer un rôle important dans l'estimation de λ. Mais avec de grands s, le premier terme ne joue plus aucun rôle, tandis que le deuxième terme s'approche de la moyenne arithmétique des y_i^*. Pour de grandes valeurs de s, le dernier terme de la formule (6.112), comme dans le cas de (6.85), donne l'estimation de μ sous la forme de la moyenne arithmétique des différences $x_i - u_i$.

Lorsque la variance σ_h^2 du bruit h_s^* dans la chaîne H^* tend vers zéro, le premier terme de l'expression (6.112) disparaît, alors que le deuxième prend la forme de la moyenne arithmétique de la constante $y_i^* = x_s^* = \lambda$, devenant ainsi une constante x^*. Dans ce cas la formule (6.112) se transforme en (6.85).

§ 3. Exemples des systèmes irréductibles

Jusque-là nous avons examiné les exemples de systèmes réductibles qui, naturellement, étaient neutres. Or, il existe une classe importante de systèmes qui ne sont pas neutres. Les systèmes de ce type sont certainement irréductibles. En effet, si un système est réductible, d'après ce qui précède, on peut le remplacer par un système en boucle ouverte; or, un système en boucle ouverte est neutre.

Par conséquent, il suffit de démontrer que le système n'est pas neutre pour prouver par là même qu'il est irréductible.

Les exemples les plus simples révèlent sans peine qu'il existe réellement des systèmes non neutres et par conséquent irréductibles. Considérons, par exemple, un objet B sans inertie dont la grandeur de sortie x_s est associée à la grandeur d'entrée v_s par la relation (fig. 6.6,*a*)

$$x_s = \mu_1 v_s + \mu_2, \quad (6.113)$$

μ_1 et μ_2 étant les variables aléatoires de densité de probabilité a priori conjointe donnée. Soit la fonction de pertes élémentaire de la forme

$$W_s = (x_s - x_s^*)^2 \quad (6.114)$$

et $x_s^* \neq 0$. Pour choisir la meilleure valeur de v_s il faut connaître alors les grandeurs μ_1 et μ_2. L'examen de la figure 6.6,*a* montre que les suites des valeurs de v_s peuvent être meilleures ou pires suivant qu'elles assurent une bonne ou une mauvaise étude du système. Si, par exemple, pour différents s on choisit la même valeur $v_s = v^1$, les renseignements obtenus ne portent que sur un point M de la caractéristique de l'objet et sont insuffisants pour définir sa pente. Mais en variant les valeurs choisies de v_s, on peut définir en principe complètement μ_1 et μ_2.

Soient les valeurs admissibles de v_s qui se situent dans l'intervalle $|v_s| \leqslant 1$ et soit x_s mesuré avec une erreur aléatoire h_s. Dans ce cas le choix des valeurs voisines v^1 et v^2 est médiocre du fait que l'erreur

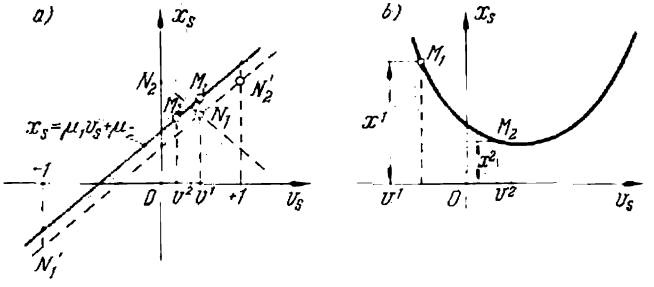


Fig. 6.6.

dans la mesure des ordonnées des points M_1 et M_2 peut entraîner une erreur importante dans la définition des paramètres μ_1 et μ_2 (pointillé passant par les points N_1 et N_2). Dans le cas considéré l'intuition suggère que les points v^1 et v^2 soient espacés le plus possible, donc placés aux extrémités de la marge admissible. Les mêmes erreurs de mesure des valeurs de x_s entraîneront alors des erreurs plus faibles dans la définition des paramètres (pointillé $N_1'N_2'$) et le même nombre de pas assurera une meilleure étude de l'objet.

Une situation analogue a lieu pour une caractéristique du type

$$x_s = \mu_1 v_s^2 + \mu_2 v_s + \mu_3. \tag{6.115}$$

Cette caractéristique est représentée sur la figure 6.6,*b*.

Dans le cas considéré, même la connaissance précise de deux points M_1 et M_2 de la caractéristique ne la définit pas complètement. Supposons qu'il faille établir le minimum de x_s; la fonction de pertes élémentaire W_s s'écrit

$$W_s = x_s. \tag{6.116}$$

La coordonnée v_s du minimum de x_s avec $\mu_1 > 0$ s'écrit

$$v_s^* = -\frac{\mu_2}{2\mu_1}. \tag{6.117}$$

Pourtant, pour définir μ_1 et μ_2, il faut résoudre un système de trois équations linéaires (deux équations ne suffisent pas, car après l'élimination de μ_3 on n'obtient qu'une équation) fournies par trois mesures qui donnent les abscisses v^1, v^2, v^3 et les ordonnées x^1, x^2, x^3 :

$$\left.\begin{aligned} x^1 &= \mu_1 (v^1)^2 + \mu_2 v^1 + \mu_3, \\ x^2 &= \mu_1 (v^2)^2 + \mu_2 v^2 + \mu_3, \\ x^3 &= \mu_1 (v^3)^2 + \mu_2 v^3 + \mu_3. \end{aligned}\right\} \tag{6.118}$$

Par conséquent, la stratégie composée uniquement de valeurs égales de v_s ou même comportant plusieurs fois l'une des deux valeurs possibles de v_s sera mauvaise. L'étude du système ne peut être

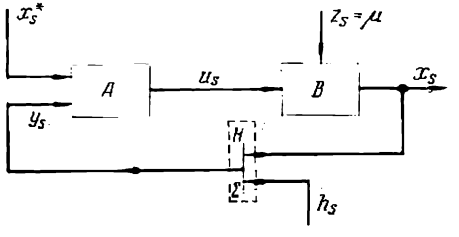


Fig. 6.7.

bonne que si le nombre de valeurs possibles de v_s est supérieur à deux et si ces valeurs sont réparties de la meilleure façon. Ce simple raisonnement entraîne déjà que dans le cas considéré il ne suffit pas de donner à v_s les valeurs extrêmes de la marge admissible.

Considérons maintenant le schéma 6.7. Soit l'équation du système B de la forme

$$x_s = \mu u_s, \tag{6.119}$$

μ étant la variable normale. La chaîne de réaction passe par le canal H avec un bruit h_s. L'équation de ce canal est

$$y_s = x_s + h_s. \tag{6.120}$$

Soit h_s la suite des variables indépendantes normales de même densité de probabilité

$$P(h_s) = q(h_s) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{h_s^2}{2\sigma_h^2}\right\}. \tag{6.121}$$

La denisité de probabilité a priori de μ s'exprime par la formule

$$P(\mu) = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right\}. \tag{6.122}$$

Soit la fonction de pertes élémentaire

$$W_s = (x_s - x_s^*)^2, \tag{6.123}$$

x_s^* étant la suite donnée des variables ($s = 0, 1, \ldots, n$). Il faut chercher la stratégie optimale de l'organe de commande A minimisant le risque R:

$$R = M\left\{\sum_{s=0}^{n} W_s\right\} = \sum_{s=0}^{n} R_s. \tag{6.124}$$

Pour résoudre ce problème, il faut déterminer la fonction α_k d'après la formule (6.58). Puisque $x_k = \mu u_k$, alors, après cette substitution, on obtient α_k sous la forme d'une intégrale par rapport seulement à μ. Etant donné que $P(x_k|\mu, u_k) = \delta(x_k - \mu u_k)$, il vient

$$\alpha_k = \alpha_k(\vec{u}_k, \vec{y}_{k-1}) = \int_{\mu=-\infty}^{\infty} (\mu u_k - x_k^*)^2 P(\mu) \times$$
$$\times \left[\prod_{i=0}^{k-1} \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{(y_i - \mu u_i)^2}{2\sigma_h^2}\right\}\right] d\mu. \tag{6.125}$$

En rejetant le facteur constant qui ne joue aucun rôle et en désignant la proportionnalité par le symbole $\equiv$, puis en remplaçant $P(\mu)$ d'après (6.122), on aboutit à l'expression

$$\alpha_k \equiv \exp\{-C_{k-1}\} \int_{-\infty}^{\infty} (\mu u_k - x_k^*)^2 \exp\{-A_{k-1}\mu^2 + B_{k-1}\mu\}\, d\mu, \tag{6.126}$$

où

$$\left.\begin{aligned} A_{k-1} &= \frac{1}{2\sigma_\mu^2} + \frac{\sum_{i=0}^{k-1} u_i^2}{2\sigma_h^2} > 0, \\ B_{k-1} &= \frac{\mu_0}{\sigma_\mu^2} + \frac{1}{\sigma_h^2} \sum_{i=0}^{k-1} u_i y_i, \\ C_{k-1} &= \frac{\mu_0^2}{2\sigma_\mu^2} + \frac{1}{2\sigma_h^2} \sum_{i=0}^{k-1} y_i^2 > 0. \end{aligned}\right\} \tag{6.127}$$

Posons

$$\mu u_k - x_k^* = z. \tag{6.128}$$

L'intégrale de (6.126) se met alors sous la forme

$$I_k = \frac{1}{u_k} \exp\left\{-A_{k-1}\left(\frac{x_k^*}{u_k}\right)^2 + B_{k-1}\left(\frac{x_k^*}{u_k}\right)\right\} \times$$
$$\times \int_{-\infty}^{\infty} z^2 \exp\left\{-\frac{A_{k-1}}{u_k^2} z^2 + \left(\frac{B_{k-1}}{u_k} - \frac{2A_{k-1}x_k^*}{u_k^2}\right) z\right\} dz. \quad (6.129)$$

La formule (6.70) rend possible le calcul de cette intégrale sous une forme explicite. En portant l'intégrale I_k dans (6.126), on obtient après des transformations simples

$$\alpha_k \equiv \exp\left\{\frac{\sigma_\mu^2 D_{k-1} - \sigma_h^2 E_{k-1}}{4\sigma_\mu^2 \sigma_h^4 A_{k-1}}\right\} \frac{1}{2A_{k-1}} \sqrt{\frac{\pi}{A_{k-1}}} \times$$
$$\times \left[u_k^2 + \frac{(2A_{k-1}x_k^* - 2B_{k-1}u_k)^2}{2A_{k-1}}\right]. \quad (6.130)$$

Nous avons posé ici

$$D_{k-1} = \left(\sum_{i=0}^{k-1} y_i u_i\right)^2 - \left(\sum_{i=0}^{k-1} u_i^2\right) \cdot \left(\sum_{i=0}^{k-1} y_i^2\right),$$
$$E_{k-1} = \mu_0^2 \sum_{i=0}^{k-1} u_i^2 - 2\mu_0 \sum_{i=0}^{k-1} y_i u_i + \sum_{i=0}^{k-1} y_i^2. \quad (6.131)$$

La première étape de définition de la stratégie optimale, la recherche de u_n^*, ne présente pas de difficulté. Le dernier facteur de (6.130) (entre crochets) étant le seul à dépendre sous une forme implicite de u_n, $\gamma_n^* = \alpha_n^* = (\alpha_n)_{\min}$ s'obtient en minimisant l'expression

$$u_n^2 + \frac{(2A_{n-1}x_n^* - 2B_{n-1}u_n)^2}{2A_{n-1}}. \quad (6.132)$$

En égalant à zéro la dérivée de cette expression par rapport à u_n, on trouve la valeur optimale

$$u_n^* = \frac{2A_{n-1}B_{n-1}x_n^*}{A_{n-1} + 2B_{n-1}^2}. \quad (6.133)$$

En portant cette valeur dans (6.130) pour $k = n$, on obtient la valeur optimale de α_n notée α_n^* ou γ_n^*:

$$\gamma_n^* = \alpha_n^* = \min_{u_n} \alpha_n = \frac{(x_n^*)^2 \sqrt{\pi \cdot A_{n-1}}}{(A_{n-1} + 2B_{n-1}^2)} \exp\left\{\frac{\sigma_\mu^2 D_{n-1} - \sigma_h^2 E_{n-1}}{4\sigma_\mu^2 \sigma_h^4 A_{n-1}}\right\}. \quad (6.134)$$

L'intégration de cette expression par rapport à y_{n-1} dans l'intervalle $-\infty < y_{n-1} < \infty$ donne le deuxième terme de la formule de γ_{n-1} (le premier terme en est α_{n-1}). Il ressort de (6.127) que A_{n-1} ne dépend pas de y_{n-1}. Mais B_{n-1} dépend de y_{n-1} de même que

D_{n-1} et E_{n-1}, cette dernière dépendance résultant de (6.131). La confrontation de (6.134) avec (6.131) et (6.127) révèle les particularités suivantes.

a) D_{n-1} et E_{n-1} sont des fonctions quadratiques de y_{n-1}; B_{n-1} dépend linéairement de y_{n-1}. Il s'ensuit que l'expression de γ_n^* est une fonction complexe de y_{n-1}. Il semble que cette intégrale ne se prête pas en général à la mise sous la forme d'une combinaison finie des paramètres. Cette difficulté est en principe surmontable si l'on fait appel à l'une des formules d'intégration approchée. Mais la solution approchée impose des calculs bien plus nombreux.

b) Plus important est le fait que A_{n-1} de même que B_{n-1}, D_{n-1} et E_{n-1} dépendent de u_{n-1} qui est présent non pas sous la forme de la différence $y_{n-1} - u_{n-1}$ (s'il en était ainsi, la dépendance de u_{n-1} pourrait être supprimée par la substitution $z = y_{n-1} - u_{n-1}$). Ainsi le résultat de l'intégration dépend de u_{n-1}. Par conséquent, le système n'est pas neutre.

La résolution du problème est rendue beaucoup plus difficile par l'impossibilité de réaliser sous une forme générale les opérations d'intégration et de minimisation. A première vue il semble que ce problème est irrésolvable même sur des calculateurs numériques universels modernes. En effet, γ_n^* est une fonction de $\vec{u}_{n-1}$ et $\vec{y}_{n-1}$, alors que l'intégrale $\int \gamma_n^* \, dy_{n-1}$ est une fonction de $\vec{u}_{n-1}$ et $\vec{y}_{n-2}$ (sans compter x_n^*), c'est-à-dire une fonction de $(2n-1)$ variables. Avec de grandes valeurs de n, la mémorisation et la minimisation de la fonction d'un aussi grand nombre de variables peuvent s'avérer pratiquement impossibles, même sur les machines les plus puissantes. Toutefois, dans de nombreux cas, on peut organiser le calcul de façon à mémoriser des fonctions d'un petit nombre de variables. Ces fonctions s'appellent parfois coordonnées, ou estimations, suffisantes ou exhaustives. Cette appellation est employée dans les mathématiques statistiques [6.8 à 6.10]. Le nombre de variables dont dépendent les estimations suffisantes n'est pas une fonction de n.

Cherchons les estimations suffisantes pour le problème considéré. Posons

$$L_k = \sum_{i=0}^{k} u_i^2, \quad M_k = \sum_{i=0}^{k} u_i y_i \quad (k = 0, 1, \ldots, n-1). \tag{6.135}$$

Nous allons montrer dans ce qui suit que les grandeurs L_{k-1} et M_{k-1} de notre exemple sont des estimations suffisantes. Elles jouissent de la propriété suivante:

$$\left.\begin{aligned} L_k &= L_{k-1} + u_k^2, \\ M_k &= M_{k-1} + u_k y_k. \end{aligned}\right\} \tag{6.136}$$

Les grandeurs A_{k-1} et B_{k-1} s'expriment aisément à l'aide de L_{k-1} et M_{k-1} :

$$\left.\begin{aligned} A_{k-1} &= \frac{1}{2\sigma_\mu^2} + \frac{L_{k-1}}{2\sigma_h^2}, \\ B_{k-1} &= \frac{\mu_0}{\sigma_\mu^2} + \frac{M_{k-1}}{\sigma_h^2}. \end{aligned}\right\} \tag{6.137}$$

C'est pourquoi, conformément à (6.126), α_k peut également être représenté sous la forme d'une fonction des variables L_{k-1}, M_{k-1}, C_{k-1}, u_k (les autres grandeurs en tant que paramètres connus ne s'écrivent pas comme arguments). L'intégration de (6.126) par rapport à μ conduit à la forme suivante de α_k :

$$\begin{aligned} \alpha_k = \exp\{-C_{k-1}\}\,[u_k^2\varphi_{2k}(L_{k-1}, M_{k-1}) - \\ - u_k x_k^* \varphi_{1k}(L_{k-1}, M_{k-1}) + (x_k^*)^2\,\varphi_{0k}(L_{k-1}, M_{k-1})], \end{aligned} \tag{6.138}$$

où

$$\left.\begin{aligned} \varphi_{2k}(L_{k-1}, M_{k-1}) &= \int_{-\infty}^{\infty} \mu^2 \exp\{-A_{k-1}\mu^2 + B_{k-1}\mu\}\,d\mu, \\ \varphi_{1k}(L_{k-1}, M_{k-1}) &= \int_{-\infty}^{\infty} \mu \exp\{-A_{k-1}\mu^2 + B_{k-1}\mu\}\,d\mu, \\ \varphi_{0k}(L_{k-1}, M_{k-1}) &= \int_{-\infty}^{\infty} \exp\{-A_{k-1}\mu^2 + B_{k-1}\mu\}\,d\mu. \end{aligned}\right\} \tag{6.139}$$

Compte tenu de (6.127) pour C_{k-1} (6.138) entraîne

$$\begin{aligned} \alpha_k \equiv \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=0}^{k-1} y_i^2\right\}\,[u_k^2\varphi_{2k}(L_{k-1}, M_{k-1}) - \\ - u_k x_k^* \varphi_{1k}(L_{k-1}, M_{k+1}) + (x_k^*)^2\,\varphi_{0k}(L_{k-1}, M_{k-1})], \end{aligned} \tag{6.140}$$

où le signe $\equiv$ indique la proportionnalité.

Maintenant on peut procéder à la définition de la stratégie optimale en ne retenant à chaque pas que les fonctions de deux variables. En effet, écrivons d'abord α_n sous la forme

$$\begin{aligned} \alpha_n \equiv \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=0}^{n-1} y_i^2\right\}\,[u_n^2\varphi_{2n}(L_{n-1}, M_{n-1}) - \\ - u_n x_n^* \varphi_{1n}(L_{n-1}, M_{n-1}) + (x_n^*)^2\,\varphi_{0n}(L_{n-1}, M_{n-1})] = \\ = \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=0}^{n-1} y_h^2\right\}\,\psi(u_n, L_{n-1}, M_{n-1}). \end{aligned} \tag{6.141}$$

L'expression entre crochets est désignée ici par $\psi(u_n, L_{n-1}, M_{n-1})$. Le paramètre x_n^* n'est pas présent ici sous une forme explicite. Minimisant α_n par rapport à u_n on trouve u_n^* et α_n^*:

$$u_n^* = u_n^*(L_{n-1}, M_{n-1}) \tag{6.142}$$

et

$$\alpha_n^* = \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=0}^{n-1} y_i^2\right\}\psi^*(L_{n-1}, M_{n-1}), \tag{6.143}$$

où

$$\psi^*(L_{n-1}, M_{n-1}) = \min_{u_n \in \Omega(u_n)} \psi(L_{n-1}, M_{n-1}, u_n). \tag{6.144}$$

On cherche ensuite la fonction

$$\begin{aligned}
\gamma_{n-1} &= \alpha_{n-1} + \int_{-\infty}^{\infty} \gamma_n^* \, dy_{n-1} = \\
&= \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=0}^{n-2} y_i^2\right\}\psi(u_{n-1}, L_{n-2}, M_{n-2}) + \\
&+ \int_{-\infty}^{\infty} \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=1}^{n-1} y_i^2\right\}\psi^*(L_{n-1}, M_{n-1})\, dy_{n-1} = \\
&= \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=0}^{n-2} y_i^2\right\}\left\{\psi(u_{n-1}, L_{n-2}, M_{n-2}) + \int_{-\infty}^{\infty} \exp\left\{\frac{1}{2\sigma_h^2} y_{n-1}^2\right\}\times\right. \\
&\left.\times\, \psi^*(L_{n-2} + u_{n-1}^2, M_{n-2} + u_{n-1}y_{n-1})\, dy_{n-1}\right\}.
\end{aligned} \tag{6.145}$$

Cette dernière transformation se fait compte tenu de (6.136). Posons

$$\begin{aligned}
\varphi(u_k, L_{k-1}, M_{k-1}) &= \psi(u_k, L_{k-1}, M_{k-1}) + \\
&+ \int_{+\infty}^{\infty} \exp\left\{\frac{1}{2\sigma_h^2} y_k^2\right\}\psi^*(L_{k-1} + u_k^2, M_{k-1} + u_k y_k)\, dy_k.
\end{aligned} \tag{6.146}$$

Si la fonction de deux variables ψ^* est connue, une méthode quelconque d'intégration approchée permet de définir l'intégrale du second membre de (6.146) aussi sous la forme d'une fonction de deux variables. ψ étant une simple combinaison des fonctions de deux variables, la fonction φ est définie en donnant les fonctions de deux variables. Ainsi

$$\gamma_{n-1} = \exp\left\{\frac{1}{2\sigma_h^2}\sum_{i=0}^{n-2} y_i^2\right\}\varphi(u_{n-1}, L_{n-2}, M_{n-2}). \tag{6.147}$$

Minimisons γ_{n-1} par rapport à u_{n-1}. Posons

$$\varphi^*(L_{k-1}, M_{k-1}) = \min_{u_k \in \Omega(u_k)} \varphi(u_k, L_{k-1}, M_{k-1}). \quad (6.148)$$

Il est évident que la minimisation de φ donne u^*_{n-1} sous la forme

$$u^*_{n-1} = u^*_{n-1}(L_{n-2}, M_{n-2}). \quad (6.149)$$

Une procédure analogue permet de définir, également par minimisation de $\varphi(u_{n-k}, L_{n-k-1}, M_{n-k-1})$, tout

$$u^*_{n-k} = u^*_{n-k}(L_{n-k-1}, M_{n-k-1}). \quad (6.150)$$

Notons que tout au long de la procédure les fonctions à retenir ne sont que celles de deux variables.

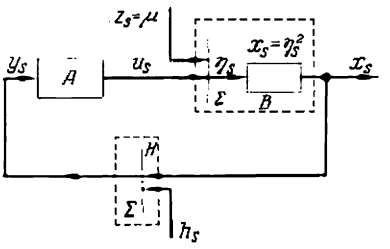


Fig. 6.8.

Considérons encore un exemple d'un système sans inertie à caractéristique quadratique (fig. 6.8). Ses équations s'écrivent

$$\left.\begin{aligned} x_s &= \eta_s^2 = (u_s + \mu)^2, \\ y_s &= x_s + h_s. \end{aligned}\right\} \quad (6.151)$$

μ est ici une variable aléatoire de densité de probabilité a priori $P_0(\mu)$. Le bruit h_s suit une loi normale de moyenne nulle et de variance σ_h^2. La grandeur x_s^* n'est pas présente sous une forme explicite (on peut poser, par exemple, $x_s^* = 0$). Toutes les valeurs de W_s sont nulles, sauf la dernière:

$$W_n = x_n. \quad (6.152)$$

Supposons que la fonction $P_0(\mu)$ ne diffère de zéro que pour $|\mu| \leqslant 1$. Posons également d'une manière analogue que les valeurs admissibles sont restreintes par la condition $|u_s| \leqslant 1$.

Le problème consiste à définir l'algorithme optimal de l'organe de commande qui vérifie la condition:

$$R = M\{W_n\} = M\{x_n\} = \min. \quad (6.153)$$

Expliquons ce problème. Supposons qu'il existe un système à caractéristique parabolique

$$x = (u + \mu)^2, \tag{6.154}$$

où μ est inconnu et la grandeur de sortie x est mesurée avec une erreur aléatoire h_s. Supposons d'abord qu'avec $i = 0, 1, \ldots, n-1$ on essaie les valeurs d'entrée u_i et on mesure les grandeurs correspondantes $y_i = x_i + h_i$. Ensuite, avec $i = n$, on affiche la valeur u_n telle qu'elle minimise l'espérance mathématique de la sortie x_n du système. En résolvant ce problème on obtient la méthode optimale pour minimiser la fonction parabolique du type (6.154) en $n+1$ pas, $t = 0, 1, \ldots, n$. Cette méthode peut être mise à la base d'un système optimal de recherche automatique.

La densité de probabilité du bruit s'exprimant par la formule

$$P(h_i) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{h_i^2}{2\sigma_h^2}\right\}, \tag{6.155}$$

il vient

$$P(y_i \mid \mu, u_i) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{(y_i - x_i)^2}{2\sigma_h^2}\right\} =$$

$$= \frac{1}{\sigma_h \sqrt{2\pi}} \exp\{a_i + b_i\mu + c_i\mu^2 + d_i\mu^3 + \mu^4\}, \tag{6.156}$$

avec

$$\left.\begin{aligned} a_i &= -\frac{(y_i - u_i^2)^2}{2\sigma_h^2}, \qquad b_i = -\frac{2u_i(u_i^2 - y_i)}{\sigma_h^2}, \\ c_i &= -\frac{1}{\sigma_h^2}(3u_i^2 - y_i), \quad d_i = -\frac{2u_i}{\sigma_h^2}. \end{aligned}\right\} \tag{6.157}$$

Tout W_k étant nul pour $k < n$, tout α_k s'annule également pour $k < n$. Il faut obtenir seulement, d'après la formule (6.58), la fonction de α_n. Compte tenu du fait que dans le problème considéré

$$P(x_k \mid \mu, u_k) = \delta\,[x_k - (u_k + \mu)^2], \tag{6.158}$$

où δ est l'impulsion-unité, après substitution $x_n = (u_n + \mu)^2$ on obtient à un facteur constant près :

$$\alpha_n \equiv \int_{-1}^{1} (u_n + \mu)^2 P_0(\mu) \exp\left\{\sum_{i=0}^{n-1} (a_i + b_i\mu + c_i\mu^2 + d_i\mu^3 + \mu^4\right\} d\mu. \tag{6.159}$$

De même que dans l'exemple précédent, on peut introduire des estimations suffisantes qui seront ici

$$\sum_{i=0}^{n-1} b_i, \quad \sum_{i=0}^{n-1} c_i, \quad \sum_{i=0}^{n-1} d_i$$

(l'exponentielle $\exp\left\{\sum_{i=0}^{n-1} a_i\right\}$ est chassée du signe somme) et dresser le schéma des calculs où ne figurent que les fonctions d'un petit nombre de variables. Pourtant, ces fonctions sont aussi assez complexes. Il est donc commode, si c'est possible, de recourir à une méthode approchée encore plus simple. Des calculs effectués par I. Timé ont montré que la fonction $b_i\mu + c_i\mu^2 + d_i\mu^3 + \mu^4$ peut être développée en série dite de Pike et de Silverberg [6.13, 6.14], c'est-à-dire mise sous la forme

$$b_i\mu + c_i\mu^2 + d_i\mu^3 + \mu^4 \cong \varphi_1(y_i, u_i) f_1(\mu) + \varphi_2(y_i, u_i) f_2(\mu). \tag{6.160}$$

Ici f_1, f_2, φ_1, φ_2 sont des fonctions qu'on peut trouver à l'aide de la procédure exposée dans [6.13, 6.14]. Si l'égalité (6.160) est approximativement vraie, de (6.159) on obtient

$$\alpha_n \cong \int_{-1}^{1} (u_n + \mu)^2 P_0(\mu) \exp\left\{A_{n-1} + \sum_{i=0}^{n-1} [\varphi_1(y_i, u_i) f_1(\mu) + \right.$$
$$\left. + \varphi_2(y_i, u_i) f_2(\mu)]\right\} d\mu = \exp\{A_{n-1}\} \int_{-1}^{1} (u_n + \mu)^2 P_0(\mu) \times$$
$$\times \exp\{E_{n-1} f_1(\mu) + F_{n-1} f_2(\mu)\} d\mu, \tag{6.161}$$

où

$$\left.\begin{aligned} A_s &= \sum_{i=0} a_i = A_{s-1} + a_s, \\ E_s &= \sum_{i=0}^{s} \varphi_1(y_i, u_i) = E_{s-1} + \varphi_1(y_s, u_s), \\ F_s &= \sum_{i=0}^{s} \varphi_2(y_i, u_i) = F_{s-1} + \varphi_2(y_s, u_s). \end{aligned}\right\} \tag{6.162}$$

L'intégration approchée de (6.161) par rapport à μ donne la fonction

$$\gamma_n^* = \alpha_n^* = \min_{u_n \in \Omega(u_n)} \alpha_n = \exp\{A_{n-1}\} \theta_n^*(E_{n-1}, F_{n-1}), \tag{6.163}$$

où θ_n^* est la fonction de deux variables E_{n-1} et F_{n-1}. La minimisation de α_n^* par rapport à u_n conduit à la commande optimale

$$u_n^* = u_n^*(E_{n-1}, F_{n-1}). \tag{6.164}$$

Cherchons ensuite la fonction

$$\gamma_{n-1} = \int_{-\infty}^{\infty} \gamma_n^* dy_{n-1} = \exp\{A_{n-2}\} \int_{-\infty}^{\infty} \exp\{a_{n-1}\} \theta_n^* [E_{n-2} +$$
$$+ \varphi_1(y_{n-1}, u_{n-1}), F_{n-2} + \varphi_2(y_{n-1}, u_{n-1})] dy_{n-1} =$$
$$= \exp\{A_{n-2}\} \theta_{n-1}(E_{n-2}, F_{n-2}, u_{n-1}), \tag{6.165}$$

où θ_{n-1} est le résultat de l'intégration qui est une fonction dépendant de trois variables E_{n-2}, F_{n-2}, u_{n-1}. Posons

$$\theta^*_{n-1} = \theta^*_{n-1}(E_{n-2}, F_{n-2}) = \min_{u_{n-1} \in \Omega(u_{n-1})} \theta_{n-1}. \tag{6.166}$$

En minimisant θ_{n-1} par rapport à u_{n-1} on trouve la commande optimale

$$u^*_{n-1} = u^*_{n-1}(E_{n-2}, F_{n-2}) \tag{6.167}$$

et la fonction θ^*_{n-1}. De plus,

$$\gamma^*_{n-1} = \exp\{A_{n-2}\}\, \theta^*_{n-1}(E_{n-2}, F_{n-2}). \tag{6.168}$$

D'une façon analogue on a

$$\gamma^*_{n-k} = \exp\{A_{n-k}\}\, \theta^*_{n-k}(E_{n-k-1}, F_{n-k-1}) \tag{6.169}$$

$(k = n, n-1, \ldots, 0)$ et les commandes optimales respectives

$$u^*_{n-k} = u^*_{n-k}(E_{n-k-1}, F_{n-k-1}). \tag{6.170}$$

La minimisation par rapport à u_{n-k} pouvant être remplacée par une comparaison de la série des intégrales $\int_{-\infty}^{\infty} \gamma^*_{n-k+1}\, dy_{n-k}$ relatives à des valeurs différentes de u_{n-k}, il en résulte qu'il ne faut retenir à chaque étape que les fonctions θ^*_{n-k}, ainsi que les fonctions obtenues u^*_{n-k}, c'est-à-dire seulement les fonctions de deux variables. L'obtention et la mémorisation de ces fonctions sont tout à fait possibles sur les calculateurs numériques actuels.

Il y a lieu de noter que la division des systèmes en réductibles et en irréductibles n'est pas basée sur la linéarité ou la non-linéarité de la caractéristique, mais dépend de l'inclusion additive ou non additive des coordonnées μ_i dans la formule de l'opérateur du système. Nous avons examiné un système réductible à objet linéaire (6.62) et un système irréductible à objet linéaire (6.119). Notre dernier exemple portait sur un système irréductible à objet non linéaire ; il n'est pas difficile non plus de donner un exemple d'un système réductible à objet non linéaire. Ainsi supposons que l'équation de l'objet B soit de la forme

$$x_i = F(\mu + v_i), \tag{6.171}$$

F étant une caractéristique monotone non linéaire. On peut alors construire une fonction inverse $\varphi = F^{-1}$. Soient les grandeurs mesurées x_i et $v_i = u_i + g_i$. Etant donné que

$$\mu + g_i = \varphi(x_i) - u_i, \tag{6.172}$$

la mesure de x_i et u_i est équivalente à la mesure de μ avec une erreur aléatoire g_i. Ainsi, un système de commande en boucle fermée peut être ramené à un système en boucle ouverte équivalent dans lequel

le résultat de la mesure de la perturbation μ passe par la chaîne E avec une erreur $e_s = g_s$.

Voici maintenant une formulation précise du concept de systèmes réductibles sans inertie. Comparons deux schémas représentés sur les figures 6.9,*a* et *b* dont l'un est en boucle fermée et l'autre en boucle ouverte.

La recherche de la stratégie optimale de l'organe de commande A du schéma 6.9,*a* peut se faire à l'aide des formules de la théorie de la commande duale. Quant à la stratégie optimale de l'organe de commande A' du schéma 6.9,*b*, elle s'obtient par des méthodes bien plus simples décrites dans le chapitre précédent. La question se pose à savoir s'il n'est pas possible de remplacer le schéma 6.9, *a* par celui de 6.9,*b*? Il se trouve qu'une telle substitution est possible pour une certaine classe de systèmes, dits réductibles. Pour donner une définition exacte des systèmes de ce type,

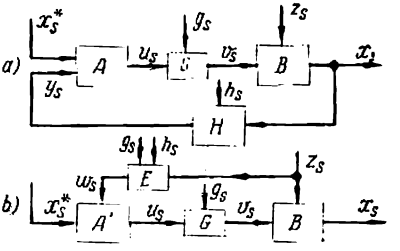


Fig. 6.9.

supposons que toutes les actions extérieures et les opérateurs G et B soient identiques pour les schémas 6.9,*a* et *b* et les systèmes soient optimaux. Si l'on peut trouver pour le schéma 6.9,*b* une fonction

$$w_s = w_s(g_s, h_s, z_s), \tag{6.173}$$

telle que les sorties u_s des organes de commande optimaux A et A' des deux systèmes soient identiques, alors le schéma 6.9,*a* est dit réductible au schéma 6.9,*b*. Les algorithmes de A et de A' peuvent différer ne serait-ce que parce que les grandeurs appliquées aux entrées de ces organes sont différentes.

A première vue on pourrait croire qu'il est possible de concevoir pour n'importe quel cas un schéma équivalent en boucle ouverte. Ainsi, pour le schéma 6.7 et les équations (6.119) et (6.120) de l'objet B et de la chaîne H respectivement on peut raisonner de la façon suivante: le fait que ces équations entraînent

$$y_s = \mu u_s + h_s \tag{6.174}$$

indique qu'on peut construire l'équivalent de y_s comme c'est représenté sur la figure 6.10. Cette grandeur peut s'obtenir en multipliant $z_s = \mu$ par u_s dans le multiplieur M, puis en additionnant les produits μu_s et h_s dans le sommateur Σ. Toutefois, on voit aisément que le schéma donné par la figure 6.10 n'est nullement identique à celui de la figure 6.9,*b*. Le schéma 6.10 prévoit une réaction locale de u_s vers y_s qui n'existe pas sur la figure 6.9, *b*. Si l'on imagine les éléments M et Σ « incorporés » dans l'organe de commande A (en for-

mant ainsi un nouvel organe A'), tout se passe alors comme si μ était simplement appliqué à l'entrée de A'. Mais un tel organe ne sera pas optimal. En effet, si la grandeur μ était simplement fournie à l'entrée de A', on pourrait alors construire un système qui afficherait aussitôt une valeur idéale de u_s. L'organe de commande A', obtenu par « incorporation » de M et Σ dans l'organe A, ne peut être

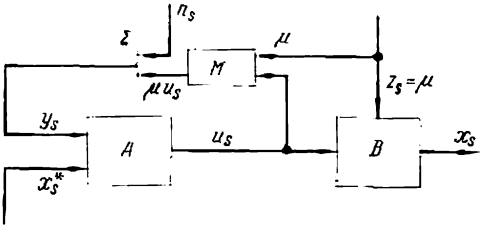


Fig. 6.10.

optimal ne serait-ce que du fait qu'à l'intérieur agit une variable aléatoire h_s. Or, nous avons déjà montré que pour cette classe de systèmes la stratégie optimale est pure.

Le schéma de la figure 6.10 se distingue en principe de celui de la figure 6.9,*b*. En effet, sur le schéma 6.10 on voit directement l'influence des valeurs de u_s sur le processus de la mesure de la perturbation $z_s = \mu$. Suivant telle ou telle grandeur de u_s, la perturbation μ est mesurée avec une précision relative plus ou moins grande. Plus u_s est grand, plus μu_s est grand par rapport au bruit h_s; par conséquent, plus la mesure de la valeur de μu_s et, donc, de la grandeur μ est précise.

Un raisonnement analogue peut être repris pour le schéma 6.8 d'équations (6.151). Dans ce cas

$$y_s = (u_s + \mu)^2 + h_s. \tag{6.175}$$

Ce qui vient d'être dit ne veut point dire que tous les systèmes neutres sont réductibles. Ainsi le système décrit par les équations (6.119) et (6.120) est irréductible, mais, lorsque les valeurs admissibles sont $u = \pm 1$, il devient neutre. Dans cet exemple, le calcul de u_s^* peut conduire à une formule définitive.

§ 4. Généralisation aux systèmes avec inertie

L'énoncé donné dans ce qui précède du problème de la théorie de la commande duale et les méthodes de sa résolution peuvent donner lieu à des généralisations différentes, dont les plus importantes sont :

a) Généralisation aux systèmes à plusieurs sorties de l'objet B admettant également plusieurs consignes.

b) Généralisation aux systèmes de commande à mémoire, c'est-à-dire à la commande des systèmes dynamiques.

c) Généralisation à une classe plus large de perturbations z_i et de consignes x_i^*. Les vecteurs des paramètres aléatoires $\bar{\mu}$ et $\bar{\lambda}$, par exemple, entrant dans les expressions de z_i et x_i respectivement, peuvent être remplacés par des processus markoviens $\bar{\mu}_i$ et $\bar{\lambda}_i$.

d) Généralisation aux systèmes à temps discret n non fixé à l'avance. On peut, par exemple, explorer le problème qui consiste à rechercher la stratégie optimale d'un organe de commande A minimisant l'espérance mathématique $M\{n\}$ du nombre de pas sous la

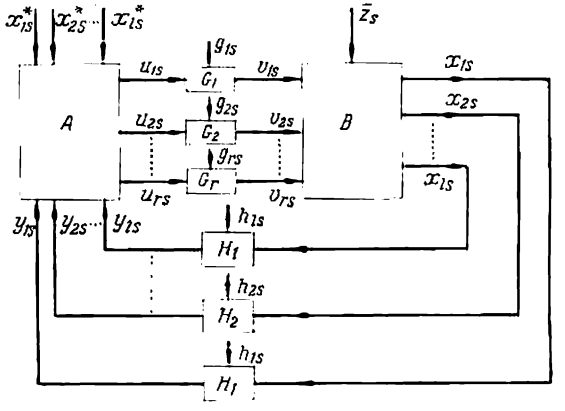


Fig. 6.11.

condition que le risque élémentaire R_n soit une grandeur suffisamment petite $R_n \leqslant \varepsilon$, où ε est une grandeur donnée qui pourtant ne doit pas être trop petite pour ne pas rendre impossible la résolution du problème. C'est un analogue du problème de commande en temps minimal.

e) Généralisation au cas des systèmes continus.

Certaines de ces généralisations sont étudiées dans ce qui suit.

Considérons la généralisation au cas des systèmes à mémoire et comportant plusieurs entrées et sorties [6.1].

Le schéma fonctionnel d'un tel système est représenté sur la figure 6.11. Posons que toutes les grandeurs sont fonctions du temps discret $t = s$ ($s = 0, 1, \ldots, n$), où n est fixé. Les commandes u_{1s}, $u_{2s}, \ldots, u_{rs}$ sont canalisées de la sortie de l'organe de commande A vers le système commandé B par les chaînes sans inertie $G_1, G_2, \ldots$ $\ldots, G_r$ dans lesquelles les signaux utiles sont entachés des bruits $g_{1s}, g_{2s}, \ldots, g_{rs}$ respectivement. De plus, l'objet B subit la perturbation $\bar{z}_s$ qui, en général, est un vecteur, c'est-à-dire un ensemble

de perturbations z_{is} ($i = 1, \ldots, w$) appliquées, peut-être, aux différentes parties du système B. Donnons le vecteur $\bar{z}_s$ de la façon suivante :

$$\bar{z}_s = \bar{z}_s(s, \bar{\mu}), \tag{6.176}$$

$\bar{\mu}$ étant le vecteur des paramètres aléatoires :

$$\bar{\mu} = (\mu_1, \mu_2, \ldots, \mu_m). \tag{6.177}$$

Les grandeurs de sortie $x_{1s}, x_{2s}, \ldots, x_{ls}$ du système B sont canalisées respectivement par les chaînes $H_1, H_2, \ldots, H_l$ vers les entrées de l'organe de commande A en s'ajoutant aux bruits $h_{1s}, h_{2s}, \ldots$ $\ldots, h_{ls}$. Admettons que les consignes $x^*_{1s}, x^*_{2s}, \ldots, x^*_{ls}$ sont données et appliquées directement aux entrées de l'organe de commande A.

Dans les cas particuliers, même lorsque le système compte plusieurs entrées, on obtient à la sortie une seule grandeur scalaire x_s ; un tel schéma se rencontre dans des systèmes d'optimisation automatique. Dans d'autres types de systèmes, la commande u_s n'est transmise à l'objet que par une chaîne ($r = 1$), alors qu'on mesure à la sortie plusieurs grandeurs x_{is} ($i = 1, \ldots, l$).

Introduisons les vecteurs composés des valeurs des variables respectives correspondant au même instant $t = s$ (vecteurs multidimensionnels ou spaciaux). Marquons ces vecteurs d'un trait sans flèche :

$$\left.\begin{array}{ll} \bar{x}^*_s = (x^*_{1s}, x^*_{2s}, \ldots, x^*_{ls}), & \bar{x}_s = (x_{1s}, x_{2s}, \ldots, x_{ls}), \\ \bar{y}_s = (y_{1s}, y_{2s}, \ldots, y_{ls}), & \bar{u}_s = (u_{1s}, u_{2s}, \ldots, u_{rs}), \\ \bar{v}_s = (v_{1s}, v_{2s}, \ldots, v_{rs}), & \bar{g}_s = (g_{1s}, g_{2s}, \ldots, g_{rs}), \\ & \bar{h}_s = (h_{1s}, h_{2s}, \ldots, h_{ls}). \end{array}\right\} \tag{6.178}$$

Il ne faut pas confondre ces vecteurs avec les vecteurs temporels affectés au-dessus d'une flèche :

$$\left.\begin{array}{ll} \vec{x}^*_{is} = (x^*_{i0}, x^*_{i1}, \ldots, x^*_{is}), & \vec{x}_{is} = (x_{i0}, x_{i1}, \ldots, x_{is}), \\ \vec{y}_{is} = (y_{i0}, y_{i1}, \ldots, y_{is}), & \vec{u}_{js} = (u_{j0}, u_{j1}, \ldots, u_{js}), \\ & \vec{v}_{js} = (v_{j0}, v_{j1}, \ldots, v_{js}) \end{array}\right\} \tag{6.179}$$

$$(i = 1, 2, \ldots, l, \; j = 1, 2, \ldots, r).$$

Avec les notations vectorielles, le schéma 6.11 peut être représenté sous une forme plus compacte (fig. 6.12). L'ensemble des chaînes $G_1, \ldots, G_r$ est remplacé ici par une seule chaîne vectorielle G et l'ensemble des chaînes $H_1, H_2, \ldots, H_l$ par une seule chaîne vectorielle H.

Imposons au type des équations du système B la contrainte suivante : supposons que ces équations aient une solution et le vecteur

$\bar{x}_s$ puisse donc être défini comme fonction du vecteur $\bar{\mu}$ et des vecteurs $\bar{v}_k$ ($k = 0, 1, \ldots, s$) :

$$\bar{x}_s = \bar{x}_s(\bar{\mu}, \bar{v}_0, \bar{v}_1, \ldots, \bar{v}_s). \tag{6.180}$$

Si les conditions initiales sont connues, on les range parmi les paramètres de la solution sans les expliciter. Mais si l'on connaît seulement la répartition probabiliste des conditions initiales, on peut les considérer alors comme coordonnées du vecteur $\bar{\mu}$.

Examinons différents cas particuliers qui peuvent se présenter pour une telle position du problème. Soit, par exemple, la trajectoire

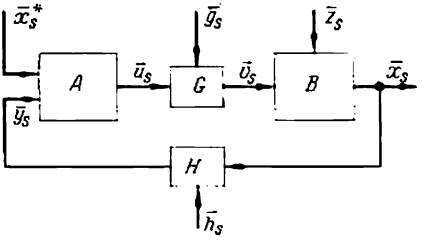


Fig. 6.12.

du système B décrite par les équations aux différences finies non linéaires :

$$x_{i,\,s+1} = F_i(s, \bar{z}_s, x_{1s}, \ldots, x_{ls}, v_{1s}, \ldots, v_{rs}) = F_i(s, \bar{\mu}, \bar{x}_s, \bar{v}_s) \tag{6.181}$$

ou

$$\bar{x}_{s+1} = \bar{F}(s, \bar{\mu}, \bar{x}_s, \bar{v}_s), \tag{6.182}$$

$\bar{F}$ étant le vecteur de coordonnées $F_1, \ldots, F_l$.

Les valeurs initiales $\bar{x}_0$ et $\bar{v}_0$ des vecteurs $\bar{x}$ et $\bar{v}$ étant imposées, cette formule permet d'obtenir le vecteur

$$\bar{x}_1 = \bar{F}(0, \bar{\mu}, \bar{x}_0, \bar{v}_0) = \Phi_1(\bar{\mu}, \bar{v}_0). \tag{6.183}$$

x_0 étant un paramètre connu, on peut l'omettre dans l'écriture générale. Connaissant $\bar{x}_1$, on peut trouver le vecteur suivant :

$$\bar{x}_2 = \bar{F}(1, \bar{\mu}, \bar{x}_1, \bar{v}_1) = \Phi_2(\bar{\mu}, \bar{v}_0, \bar{v}_1). \tag{6.184}$$

Dans cette équation $\bar{x}_1$ est remplacé par l'expression (6.183). En poursuivant cette procédure on aboutit à la formule du type (6.180) du vecteur $\bar{x}_s$.

Un autre exemple peut être fourni par le schéma fonctionnel de la figure 6.13. Les parties B_1 et B_3 de l'objet B sont linéaires de fonctions de transfert

$$x_{1s} = \sum_{k=0}^{s} a_{1k}(s, \bar{\mu})\, v_{s-k}, \quad x_{3s} = \sum_{k=0}^{s} a_{3k}(s, \bar{\mu})\, x_{2,\, s-k}. \qquad (6.185)$$

La partie B_2 est non linéaire sans inertie; son équation s'écrit

$$x_{2s} = F(s, \bar{\mu}, x_{1s}). \qquad (6.186)$$

La relation entre x_{2s} et x_{1s} doit être univoque mais pas forcément biunivoque.

Les équations (6.185) et (6.186) montrent que les grandeurs x_{1s}, x_{2s}, x_{3s} peuvent être exprimées en fonction de $\bar{\mu}$ et du vecteur $\vec{v} = (v_0, v_1, \ldots, v_s)$, c'est-à-dire qu'il est possible d'établir une relation du type (6.180).

Ainsi, dans le cas des classes très importantes et assez générales, la caractéristique du système commandé B peut être ramenée au

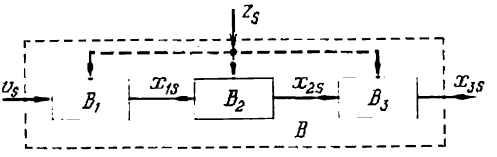


Fig. 6.13.

type (6.180). Dans ce qui suit nous allons considérer que cette réduction est faite et que la fonction (6.180) est donnée.

Introduisons des matrices spatio-temporelles que nous allons marquer d'un trait et d'une flèche placés au-dessus du symbole. Toute matrice de ce type est, d'une part, un ensemble de vecteurs temporels, c'est-à-dire un vecteur colonne dont les « coordonnées » sont des vecteurs temporels. Mais, d'autre part, cette même matrice peut être considérée comme un ensemble de vecteurs spatiaux, c'est-à-dire comme un vecteur ligne dont les «coordonnées » sont des vecteurs spatiaux. Par exemple,

$$\vec{\bar{u}}_s = \begin{vmatrix} u_{10}, & u_{11}, & \ldots, & u_{1s} \\ u_{20}, & u_{21}, & \ldots, & u_{2s} \\ \vdots & \vdots & & \\ u_{r0}, & u_{r1}, & \ldots, & u_{rs} \end{vmatrix} = \begin{vmatrix} \vec{u}_{1s} \\ \vec{u}_{2s} \\ \vdots \\ \vec{u}_{rs} \end{vmatrix} = |\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_s|. \qquad (6.187)$$

De la même façon on peut introduire d'autres matrices spatio-temporelles :

$$\vec{\bar{v}}_s = \begin{vmatrix} v_{10}, \ldots, v_{1s} \\ \vdots \\ v_{r0}, \ldots, v_{rs} \end{vmatrix} = \begin{vmatrix} \vec{v}_{1s} \\ \vdots \\ \vec{v}_{rs} \end{vmatrix} = |\bar{v}_0, \ldots, \bar{v}_s|, \tag{6.188}$$

$$\vec{\bar{x}}_s = \begin{vmatrix} x_{10}, \ldots, x_{1s} \\ \vdots \\ x_{l0}, \ldots, x_{ls} \end{vmatrix} = \begin{vmatrix} \vec{x}_{1s} \\ \vdots \\ \vec{x}_{ls} \end{vmatrix} = |\bar{x}_0, \ldots, \bar{x}_s|, \tag{6.189}$$

etc.

Choisissons la fonction de pertes du système sous la forme

$$W = \sum_{s=0}^{n} W_s, \tag{6.190}$$

où

$$W_s = W_s(s, \bar{\mu}, \bar{x}_s^*, \bar{x}_s). \tag{6.191}$$

Des cas peuvent se présenter lorsque W_s dépend explicitement de μ. Supposons, par exemple, que dans le schéma de la figure 6.12 la relation entre la sortie unique x_s et les entrées est définie par la formule

$$x_s = \mu_{r+1} + \sum_{i,\, j=1}^{r} a_{ij} (v_{is} - \mu_i)(v_{js} - \mu_j), \tag{6.192}$$

où a_{ij} = const et v_{js} sont les entrées du système. Supposons ensuite que le deuxième terme du second membre de (6.192) ne peut être négatif. La valeur minimale possible de x_s est alors μ_{r+1}, et les valeurs $v_{is} = \mu_i$ minimisent x_s. L'écart de x_s par rapport au minimum s'exprime par la différence $x_s - \mu_{r+1}$. C'est pourquoi, dans le cas considéré, si la recherche du minimum de x_s doit être automatique, on peut poser

$$W_s = x_s - \mu_{r+1} \tag{6.193}$$

et imposer la minimisation de l'espérance mathématique de la fonction W correspondante.

Supposons donnée la densité de probabilité a priori du vecteur $\bar{\mu}$ que nous allons noter $P(\bar{\mu}) = P_0(\bar{\mu})$ et les densités de probabilité $P(g_{js})$ ($j = 1, 2, \ldots, r$) des variables aléatoires g_{js}. Pour j fixé, les variables g_{js} représentent une suite de variables indépendantes ayant la même densité de probabilité $P(g_{js})$ indépendante du temps. Pour différentes chaînes (j différents) les densités de probabilité peuvent être différentes.

Les mêmes conditions sont vraies pour les densités de probabilité $P(h_{is})$ des grandeurs h_{is} $(i = 1, 2, \ldots, l)$.

Convenons que toutes les actions extérieures subies par le système sont indépendantes entre elles.

Soient les chaînes G et H sans mémoire. Il vient

$$v_{js} = v_j(g_{js}, u_{js}), \quad y_{is} = y_i(h_{is}, x_{is}). \tag{6.194}$$

Les données indiquées ci-dessus permettent d'obtenir les densités de probabilité conditionnelles $P(y_{is}|x_{is})$, $P(v_{js}|u_{js})$ et, par conséquent, $P(\bar{y}_s|\bar{x}_s)$ et $P(\bar{v}_s|\bar{u}_s)$. On peut donc établir également les densités de probabilité conditionnelles

$$\begin{aligned} P(\vec{y}_{is}|\vec{x}_{is}) &= \prod_{\nu=0}^{s} P(y_{i\nu}|x_{i\nu}), \\ P(\vec{v}_{js}|\vec{u}_{js}) &= \prod_{\nu=0}^{s} P(v_{j\nu}|u_{j\nu}). \end{aligned} \tag{6.195}$$

D'après ces densités de probabilité on peut définir également les densités de probabilité conditionnelles des matrices $P(\vec{\bar{y}}_s|\vec{\bar{x}}_s)$ et $P(\vec{\bar{v}}_s|\vec{\bar{u}}_s)$.

L'algorithme de l'organe de commande A est caractérisé par une densité de probabilité Γ_s de son vecteur de sortie $\bar{u}_s$, dépendant en général de toutes les valeurs antérieures des entrées de l'organe de commande et des valeurs courantes des consignes:

$$P_s(\bar{u}_s) = \Gamma_s(\bar{u}_s|\vec{\bar{x}}_s^*, \vec{\bar{y}}_{s-1}, \vec{\bar{u}}_{s-1}). \tag{6.196}$$

Le problème consiste à trouver une suite des fonctions Γ_i $(i = 1, \ldots, n)$ et une densité de probabilité initiale $\Gamma_0 = P(u_0)$ telles qu'elles minimisent le risque R, espérance mathématique de la fonction de pertes W définie par (6.190) et (6.191).

D'après (6.191) et (6.180)

$$W_s = W_s[s, \bar{\mu}, \bar{x}_s^*, \bar{x}_s(s, \bar{\mu}, \vec{\bar{v}})] = W_s(s, \bar{\mu}, \bar{x}_s^*, \vec{\bar{v}}). \tag{6.197}$$

Bornons-nous à déduire l'expression *) du risque R, la matrice $\vec{\bar{x}}_s^*$ étant fixée.

Supposons connue la densité de probabilité conditionnelle conjointe

$$P(\bar{\mu}, \vec{\bar{u}}_s, \vec{\bar{v}}_s, \vec{\bar{y}}_{s-1}|\vec{\bar{x}}_s^*).$$

*) Cette déduction peut se faire de la même façon que pour le cas des systèmes sans mémoire. Dans ce qui suit nous donnons une autre déduction.

Dans cette expression figure la matrice $\vec{y}_{s-1}$ du fait que d'après (6.196) la matrice $\vec{u}_s$ est associée précisément à $\vec{y}_{s-1}$.

Le risque élémentaire R_s relatif à l'instant $t = s$ s'écrit

$$R_s = M\{W_s \mid \vec{x}_s^*\} = \int\limits_{\Omega(\bar{\mu}, \vec{u}_s, \vec{v}_s, \vec{y}_{s-1})} W_s(s, \bar{\mu}, \vec{x}_s^*, \vec{v}_s) \times$$
$$\times P(\bar{\mu}, \vec{u}_s, \vec{v}_s, \vec{y}_{s-1} \mid \vec{x}_s^*)\, d\Omega. \quad (6.198)$$

Le théorème du produit des probabilités entraîne

$$P(\bar{\mu}, \vec{u}_s, \vec{v}_s, \vec{y}_{s-1} \mid \vec{x}_s^*) = P(\mu)\, P(\vec{v}_s, \vec{u}_s, \vec{y}_{s-1} \mid \bar{\mu}, \vec{x}_s^*). \quad (6.199)$$

Cette formule est vraie si seulement $\bar{\mu}$ et $\vec{x}_s^*$ sont indépendants.

Considérons le deuxième facteur du second membre de l'égalité (6.199) et procédons à sa transformation :

$$P(\vec{v}_s, \vec{u}_s, \vec{y}_{s-1} \mid \bar{\mu}, \vec{x}_s^*) = P(\vec{u}_s, \vec{y}_{s-1} \mid \bar{\mu}, \vec{x}_s^*)\, P(\vec{v}_s \mid \vec{u}_s, \vec{y}_{s-1}, \bar{\mu}, \vec{x}_s^*) =$$
$$= P(\vec{u}_s, \vec{y}_{s-1} \mid \bar{\mu}, \vec{x}_s^*)\, P(\vec{v}_s \mid \vec{u}_s, \vec{y}_{s-1}, \bar{\mu}). \quad (6.200)$$

Cette dernière transformation est légitime car la densité de probabilité de la matrice $\vec{v}_s$ est complètement définie en fixant $\vec{u}_s$, $\vec{y}_{s-1}$, $\bar{\mu}_s$ et une information supplémentaire sur $\vec{x}_s^*$ ne peut pas la modifier.

Il n'est pas difficile de montrer ensuite par analogie avec (6.25) que

$$P(\vec{u}_s, \vec{y}_{s-1} \mid \bar{\mu}, \vec{x}_s^*) = P(\bar{u}_0, \bar{y}_0 \mid \bar{\mu}, \vec{x}_0^*) \cdot P(\bar{u}_1, \bar{y}_1 \mid \bar{\mu}, \bar{u}_0, \bar{y}_0, \vec{x}_1^*) \times$$
$$\times P(\bar{u}_2, \bar{y}_2 \mid \bar{\mu}, \vec{u}_1, \vec{y}_1, \vec{x}_2^*)\; P(\bar{u}_i, \bar{y}_i \mid \bar{\mu}, \vec{u}_{i-1}, \vec{v}_{i-1}, \vec{x}_i^*) \times \ldots$$
$$\ldots \times P(\bar{u}_{s-1}, \bar{y}_{s-1} \mid \bar{\mu}, \vec{u}_{s-2}, \vec{y}_{s-2}, \vec{x}_{s-1}^*)\; P(\bar{u}_s \mid \bar{\mu}\, \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{x}_s^*). \quad (6.201)$$

Considérons l' i-ième facteur de ce produit $(0 < i < s)$:

$$P(\bar{u}_i, \bar{y}_i \mid \bar{\mu}, \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{x}_i^*) =$$
$$= P(\bar{y}_i \mid \bar{\mu}, \vec{u}_i, \vec{y}_{i-1}, \vec{x}_i^*) \cdot P(\bar{u}_i \mid \bar{\mu}, \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{x}_i^*) =$$
$$= P(\bar{y}_i \mid i, \bar{\mu}, \vec{y}_{i-1}, \vec{u}_i) \cdot \Gamma_i(\bar{u}_i \mid \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{x}_i^*). \quad (6.202)$$

En effet, la densité de probabilité de $\bar{y}_i$ s'obtient ayant fixé $\bar{\mu}$, $\vec{y}_{i-1}$ et $\vec{u}_i$ et une fixation supplémentaire de $\vec{x}_i^*$ ne la changera pas. Dans la formule de la densité de probabilité de $\bar{y}_i$ on souligne sa dépendance explicite de l'instant i.

Le deuxième facteur de (6.202) Γ_i est la densité de probabilité de $\bar{u}_i$, qui d'après (6.196) ne dépend que de $\vec{\bar{u}}_{i-1}$, $\vec{\bar{y}}_{i-1}$, $\vec{\bar{x}}_i^*$.

En portant (6.202) dans (6.201), on a :

$$P(\vec{\bar{u}}_s, \vec{\bar{y}}_{s-1} \mid \bar{\mu}, \vec{\bar{x}}_s^*) = \prod_{i=0}^{s} \Gamma_i \cdot \prod_{i=0}^{s-1} P(\bar{y}_i \mid i, \bar{\mu}_s, \vec{\bar{y}}_{i-1}, \vec{\bar{u}}_i). \quad (6.203)$$

Nous avons posé ici

$$P(\bar{y}_0 \mid i, \vec{\bar{y}}_{-1}, \bar{\mu}, \vec{\bar{u}}_0) = P(\bar{y}_0 \mid i, \bar{\mu}, u_0); \; \Gamma_0 = P_0(\bar{u}_0, \vec{\bar{x}}_0^*). \quad (6.204)$$

Portons l'expression (6.202) dans (6.200), puis (6.200) dans (6.199) et (6.199) dans (6.198). L'expression de R_s devient alors

$$R_s = \int\limits_{\Omega(\bar{\mu}, \vec{\bar{v}}_s, \vec{\bar{u}}_s, \vec{\bar{y}}_{s-1})} W[s, \bar{\mu}, \vec{\bar{x}}_s^*, \vec{\bar{v}}_s] \cdot P(\bar{\mu}) \cdot P(\vec{\bar{v}}_s \mid \vec{\bar{u}}_s, \vec{\bar{y}}_{s-1}, \bar{\mu}) \times$$
$$\times \prod_{i=0}^{s-1} P(\bar{y}_i \mid i, \vec{\bar{y}}_{i-1}, \bar{\mu}, \vec{\bar{u}}_i) \cdot \prod_{i=0}^{s} \Gamma_i(\bar{u}_i \mid \vec{\bar{u}}_{i-1}, \vec{\bar{y}}_{i-1}, \vec{\bar{x}}_i^*)\, d\Omega. \quad (6.205)$$

Le risque total R est toujours défini par l'égalité

$$R = \sum_{s=0}^{n} R_s. \quad (6.206)$$

Dans le cas particulier du système B ne comportant qu'une entrée et une sortie les matrices spatio-temporelles de la formule (6.205) sont remplacées par les vecteurs temporels et les vecteurs spatiaux, par des scalaires. Dans ces conditions, l'expression (6.205) s'écrit

$$R_s = \int\limits_{\Omega(\bar{\mu}, \vec{v}_s, \vec{u}_s, \vec{y}_{s-1})} W_s(s, \bar{\mu}, \vec{\bar{x}}_s^*, \vec{v}_s) \cdot P(\bar{\mu}) \cdot P(\vec{v}_s \mid \vec{u}_s, \vec{y}_{s-1}, \bar{\mu}) \times$$
$$\times \prod_{i=0}^{s-1} P(y_i \mid i, \bar{\mu}, \vec{y}_{i-1}, \vec{u}_i) \prod_{i=0}^{s} \Gamma_i(u_i \mid \vec{u}_{i-1}, \vec{y}_{i-1}, \vec{\bar{x}}_i^*)\, d\Omega. \quad (6.207)$$

Cette expression est plus générale que (6.57) puisqu'elle tient compte de la mémoire que possède le système B. Aussi dans la formule (6.207) W_s dépend-elle non pas du scalaire v_s, mais du vecteur temporel $\vec{v}_s$, ce qui entraîne la nécessité de prendre en considération la densité de probabilité $P(\vec{v}_s \mid \vec{u}_s, \vec{y}_{s-1}, \bar{\mu})$. De plus, la densité de probabilité conditionnelle de y_i dépend maintenant non seulement de la valeur de u_i mais aussi du vecteur $\vec{u}_i$ tout entier, c'est-à-dire de la

« préhistoire » de la grandeur de sortie du système B ainsi que de $\vec{y}_{i-1}$. En l'absence de la perturbation g_s, lorsque par exemple $u_s = v_s$, $P(y_i)$ ne dépend pas de $\vec{y}_{i-1}$.

Pour définir la stratégie optimale considérons, de même qu'au § 2 de ce chapitre, le risque élémentaire R_n, en supposant que les densités Γ_i $(i = 0, 1, \ldots, n-1)$ soient fixées. Posons

$$\alpha_k = \alpha_k(\vec{u}_k, \vec{y}_{k-1}, \bar{x}_k^*) = \int\limits_{\Omega(\bar{\mu}, \vec{v}_k)} W_k(k, \bar{\mu}, \bar{x}_k^*, \vec{v}_k) \times$$

$$\times P(\bar{\mu}) \cdot P(\vec{v}_k \mid \vec{u}_k, \vec{y}_{k-1}, \bar{\mu}) \cdot \prod_{i=0}^{k-1} P(\bar{y}_i \mid i, \vec{y}_{i-1}, \bar{\mu}, \vec{u}_i)\, d\Omega. \quad (6.208)$$

Pour $k=0$, la grandeur $\prod\limits_{i=0}^{k} P(\bar{y}_i)$ doit être égale à l'unité. Désignons également

$$\beta_k = \beta_k(\vec{u}_k, \vec{x}_k^*, \vec{y}_{k-1}) = \prod_{i=0}^{k} \Gamma_i. \quad (6.209)$$

On a

$$R_n = \int\limits_{\Omega(\vec{u}_n, \vec{y}_{n-1})} \alpha_n(\bar{u}_n, \vec{u}_{n-1}, \bar{x}_n^*) \cdot \beta_{n-1} \cdot \Gamma_n\, d\Omega =$$

$$= \int\limits_{\Omega(\vec{u}_{n-1}, \vec{y}_{n-1})} \beta_{n-1} \varkappa_n(\vec{u}_{n-1}, \vec{y}_{n-1}, \vec{x}_n^*)\, d\Omega, \quad (6.210)$$

avec

$$\varkappa_n(\vec{u}_{n-1}, \vec{y}_{n-1}, \vec{x}_n^*) =$$

$$= \int\limits_{\Omega(\bar{u}_n)} \alpha_n(\bar{u}_n, \vec{u}_{n-1}, \vec{y}_{n-1}, \bar{x}_n^*)\, \Gamma_n(\bar{u}_n, \vec{u}_{n-1}, \vec{y}_{n-1}, \vec{x}_n^*)\, d\Omega. \quad (6.211)$$

Le choix de la fonction Γ_n qui est une densité de probabilité est restreint par la condition

$$\int\limits_{\Omega(\bar{u}_n)} \Gamma_n(\bar{u}_n, \vec{u}_{n-1}, \vec{y}_{n-1}, \vec{x}_n^*)\, d\Omega = 1. \quad (6.212)$$

Le choix de la fonction Γ_n vérifiant la condition (6.212) doit assurer la minimisation de la grandeur R_n. Or, cette grandeur sera minimale si pour tous $\vec{u}_{n-1}$, $\vec{y}_{n-1}$, la grandeur $\varkappa_n$ est minimale, ce

que l'on obtient en prenant

$$\Gamma_n = \delta(\bar{u}_n - \bar{u}_n^*). \tag{6.213}$$

Ici δ est l'impulsion unité et la grandeur $\bar{u}_n^*$ est définie par la condition

$$\gamma_n^* = \alpha_n(\bar{u}_n^*, \vec{\bar{u}}_{n-1}, \vec{\bar{y}}_{n-1}, \bar{x}_n^*) = \min_{u_n \in \Omega(\bar{u}_n)} \alpha_n(\bar{u}_n, \vec{\bar{u}}_{n-1}, \vec{\bar{y}}_{n-1}, \bar{x}_n^*). \tag{6.214}$$

D'après (6.214) la commande optimale $\bar{u}_n^*$ dépend de $\vec{\bar{u}}_{n-1}$, $\vec{\bar{y}}_{n-1}$, $\bar{x}_n^*$:

$$\bar{u}_n^* = \bar{u}_n^*(\vec{\bar{u}}_{n-1}, \vec{\bar{y}}_{n-1}, \bar{x}_n^*). \tag{6.215}$$

Ainsi on établit que la stratégie optimale Γ_n n'est pas aléatoire, mais pure. Le choix de la commande optimale $\bar{u}_n^*$ est donné par les formules (6.214) et (6.215).

En passant d'une façon analogue à l'examen de la somme des termes $R_{n-1} + R_n$, puis de $R_{n-2} + R_{n-1} + R_n$, etc., on peut trouver les stratégies optimales élémentaires Γ_i $(i = n-1, n-2, \ldots)$. Une démonstration de ce type est donnée au § 1. Voici ce qui en résulte. Introduisons la fonction

$$\gamma_{n-k} = \alpha_{n-k} + \int_{\Omega(\bar{u}_{n-k})} \gamma_{n-k+1}^* \, d\Omega, \tag{6.216}$$

en outre, $\gamma_n = \alpha_n$ et

$$\gamma_{n-k}^* = (\gamma_{n-k})\, \bar{u}_{n-k} = \bar{u}_{n-k}^*, \tag{6.217}$$

alors que la valeur de $\bar{u}_{n-k}^*$ est fournie par la condition

$$\gamma_{n-k}(\bar{u}_{n-k}^*, \vec{\bar{u}}_{n-k-1}, \vec{\bar{y}}_{n-k-1}, \bar{x}_{n-k}^*) =$$
$$= \min_{\bar{u}_{n-k} \in \Omega(\bar{u}_{n-k})} \gamma_{n-k}(\bar{u}_{n-k}, \vec{\bar{u}}_{n-k-1}, \vec{\bar{y}}_{n-k-1}, \bar{x}_{n-k}^*). \tag{6.218}$$

Il est évident que

$$\bar{u}_{n-k}^* = \bar{u}_{n-k}^*(\vec{\bar{u}}_{n-k-1}, \vec{\bar{y}}_{n-k-1}, \bar{x}_{n-k}^*). \tag{6.219}$$

La stratégie optimale Γ_{n-k}^* est définie alors par l'expression

$$\Gamma_{n-k}^* = \delta(\bar{u}_{n-k} - \bar{u}_{n-k}^*), \tag{6.220}$$

c'est-à-dire c'est une stratégie pure ; de plus, la commande optimale $\bar{u}_{n-k}^*$ donnée par (6.219) dépend de la « préhistoire » des grandeurs d'entrée de l'organe de commande A ainsi que de la valeur courante de $\bar{x}_{n-k}^*$.

On généralise aisément les formules déduites ci-dessus au cas de la consigne x_s^* fournie à l'entrée de l'organe de commande par une chaîne comportant des perturbations. La déduction est analogue à celle donnée au § 1 de ce chapitre.

A titre d'exemple de l'application des formules obtenues, examinons le problème de synthèse de l'algorithme d'un organe de commande optimal A pour un système de stabilisation automatique. Le

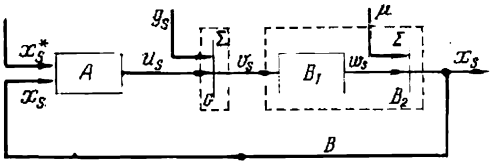


Fig. 6.14.

schéma fonctionnel du système est donné par la figure 6.14. Le système commandé B est composé de la partie B_1 munie d'une mémoire et de la partie B_2 sans mémoire. Les équations de ces parties s'écrivent

$$w_s = \sum_{k=0}^{s} a_k v_{s-k}, \quad x_s = \mu + w_s. \tag{6.221}$$

a_k sont ici les constantes données, μ la variable aléatoire de densité de probabilité a priori donnée $P(\mu) = P_0(\mu)$. L'équation générale de l'objet B peut s'écrire

$$x_s = \mu + \sum_{k=0}^{s} a_k v_{s-k}. \tag{6.222}$$

Les termes de la suite des variables indépendantes g_s ont la même densité de probabilité $q(g_s)$. Les fonctions $P(\mu)$ et $q(g_s)$ sont des densités de probabilité normales

$$\begin{aligned} P(\mu) &= \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{\mu^2}{2\sigma_\mu^2}\right\}, \\ q(g_s) &= \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{g_s^2}{2\sigma_g^2}\right\}. \end{aligned} \tag{6.223}$$

L'équation de la chaîne G est de la forme

$$v_s = g_s + u_s. \tag{6.224}$$

La fonction de pertes élémentaire est définie par l'expression

$$W_s = (x_s - x_s^*)^2. \tag{6.225}$$

Dans ces conditions la fonction de pertes totale devient

$$W = \sum_{s=0}^{n} W_s = \sum_{s=0}^{n} (x_s - x_s^*)^2. \tag{6.226}$$

Dans le cas particulier de systèmes à entrée et sortie uniques, la formule (6.208) permet d'obtenir une expression de α_k sous la forme suivante

$$\alpha_k = \alpha_k(\vec{u}_k, \vec{y}_{k-1}, x_k^*) = \\ = \int\limits_{\Omega(\bar{\mu}, \vec{v}_k)} W(k, \bar{\mu}, x_k^*, \vec{v}_k) P(\bar{\mu}) \cdot P(\vec{v}_k \mid \vec{u}_k, \bar{y}_{k-1}, \bar{\mu}) \times \\ \times \prod_{i=0}^{k-1} P(y_i \mid i, \bar{\mu}, \vec{y}_{i-1}, \vec{u}_i)\, d\Omega. \tag{6.227}$$

Dans l'exemple considéré *)

$$\alpha_k = \int\limits_{\Omega(\mu, \vec{v}_k)} \left(\mu - x_k^* + \sum_{\rho=0}^{k} \alpha_\rho v_{k-\rho}\right)^2 \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{\mu^2}{2\sigma_\mu^2}\right\} \times \\ \times P(\vec{v}_k \mid \vec{u}_k, \vec{x}_{k-1}, \bar{\mu}) \cdot \prod_{i=0}^{k-1} P(y_i \mid i, \mu, \vec{y}_{i-1}, \vec{u}_i). \tag{6.228}$$

Cherchons les expressions $P(\vec{v}_k \mid \vec{u}_k, \vec{x}_{k-1}, \bar{\mu})$ et $P(y_i \mid i, \mu, \vec{y}_{i-1}, \vec{u}_i)$.

Il ressort de la formule (6.224) que

$$P(v_i \mid u_i) = q(v_i - u_i) = \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{(v_i - u_i)^2}{2\sigma_g^2}\right\}. \tag{6.229}$$

Etant donné que les grandeurs aléatoires g_i $(i = 0, 1, \ldots, k)$ sont indépendantes, v_k ne dépend que de u_k. Par conséquent,

$$P(\vec{v}_k \mid \vec{u}_k, x_{k-1}, \bar{\mu}) = P(v_k \mid u_k) \cdot P(\vec{v}_{k-1} \mid \vec{u}_{k-1}, \vec{x}_{k-1}, \bar{\mu}) = \\ = \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{(v_k - u_k)^2}{2\sigma_g^2}\right\} \cdot P(\vec{v}_{k-1} \mid \vec{u}_{k-1}, \vec{x}_{k-1}, \bar{\mu}). \tag{6.230}$$

Cherchons d'abord $P(y_i \mid i, \mu, \vec{y}_{i-1}, \vec{u}_i) = P(y_i \mid \mu, \vec{y}_{i-1}, \vec{u}_i)$.

L'indice i à droite du trait vertical peut être omis car les équations du système B et de la chaîne G ne dépendent pas de i.

---

*) [6.1], dans cet exemple, ne prend pas en considération la dépendance de $P(y_i)$ et $P(\vec{v}_i)$ de $\vec{y}_{i-1}$, ce qui simplifie les formules mais risque d'aboutir à l'algorithme qui s'avère loin d'être optimal. Nous donnons ici des formules précises.

On tire de la formule (6.222) :

$$\left.\begin{aligned} x_i &= \mu + \sum_{\rho=0}^{i} a_\rho v_{i-\rho} = \mu + \sum_{\rho=0}^{i} a_\rho (u_{i-\rho} + g_{i-\rho}), \\ x_{i-1} &= \mu + \sum_{\rho=0}^{i-1} a_\rho v_{i-1-\rho} = \mu + \sum_{\rho=0}^{i-1} a_\rho (u_{i+\rho} + g_{i+\rho}), \\ &\dots\dots\dots\dots\dots\dots \\ x_0 &= \mu + a_0 v_0 = \mu + a_0 (u_0 + g_0). \end{aligned}\right\} \tag{6.231}$$

C'est un système d'équations linéaires par rapport à $g_0, g_1, \dots, g_i$. Récrivons-le sous la forme suivante:

$$\left.\begin{aligned} &\sum_{\rho=0}^{i} a_\rho g_{i-\rho} = x_i - \mu - \sum_{\rho=0}^{i} a_\rho u_{i-\rho}, \\ &\sum_{\rho=0}^{i} a_\rho g_{i-1-\rho} = x_{i-1} - \mu - \sum_{\rho=0}^{i-1} a_\rho u_{i-1-\rho}, \\ &\dots\dots\dots\dots\dots\dots \\ &a_0 g_0 = x_0 - \mu - a_0 u_0. \end{aligned}\right\} \tag{6.232}$$

Désignons les seconds membres de ces équations par b_k :

$$\left.\begin{aligned} &x_i - \mu - \sum_{\rho=0}^{i} a_\rho u_{i-\rho} = b_i, \\ &x_{i-1} - \mu - \sum_{\rho=0}^{i} a_\rho u_{i-1-\rho} = b_{i-1}, \\ &\dots\dots\dots\dots\dots\dots \\ &x_0 - \mu - a_0 u_0 = b_0. \end{aligned}\right\} \tag{6.233}$$

Si on élimine du système (6.232) les variables $g_0, g_1, \dots, g_{i-1}$, on obtient la relation entre x_i et g_i pour $\vec{x}_{i-1}$, μ et $\bar{u}_i$ étant fixés. Désignons le déterminant du système (6.232) par Δ :

$$\Delta = \begin{vmatrix} a_i & a_{i-1} & \dots & a_1 & a_0 \\ a_{i-1} & a_{i-2} & \dots & a_0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_0 & \dots & & 0 \\ a_0 & 0 & \dots & & 0 \end{vmatrix}. \tag{6.234}$$

Comme a_0 n'est pas nul (s'il n'en était pas ainsi l'équation (6.222) entraînerait que $x_0 = \mu$, c'est-à-dire que le paramètre μ est connu exactement), le déterminant Δ n'est non plus égal à zéro.

En résolvant le système d'équations on trouve g_i :

$$g_i = \sum_{j=0}^{i} \varkappa_{ij} b_j, \tag{6.235}$$

$\varkappa_{ij}$ étant les constantes exprimées à l'aide de $a_0, \ldots, a_i$ d'après les formules connues :

$$\varkappa_{ij} = \frac{1}{\Delta} \begin{vmatrix} a_i & a_{i-1} & \ldots & a_1 & 0 \\ a_{i-1} & a_{i-2} & \ldots & a_0 & 0 \\ \cdot & \cdot & \cdot & \cdot & 1 \\ a_1 & a_0 & \ldots & & 0 \\ a_0 & 0 & \ldots & & 0 \end{vmatrix} \quad (j\text{-ième ligne}). \tag{6.236}$$

La formule (6.233) montre que seul b_i dépend de la grandeur x_i ; cette dernière, affectée du coefficient $\varkappa_{ii}$, est présente linéairement dans l'expression (6.235).

C'est pourquoi la densité de probabilité $P(x_i \,|\, \mu, \vec{u}_i, \vec{x}_{i-1})$ peut prendre la forme suivante :

$$P(x_i \,|\, \mu, \vec{u}_i, \vec{x}_{i-1}) = |\varkappa_{ii}| \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp\left\{ -\frac{\left(\sum_{j=0}^{i} \varkappa_{ij} b_j\right)^2}{2\sigma_g^2} \right\}. \tag{6.237}$$

Maintenant on peut écrire l'expression définitive de $P(\vec{v}_{k-1} \,|\, \vec{u}_{k-1}, \vec{x}_{k-1}, \mu)$ de la formule (6.230). Les valeurs $v_0, \ldots, v_{k-1}$ s'obtiennent directement à partir de $\mu, x_0, \ldots, x_{k-1}$ données, d'après (6.231). Les solutions de ces équations s'écrivent d'une manière analogue à (6.235) mais en remplaçant g_i par v_i ; maintenant par b_j on entend les grandeurs $x_j - \mu$ $(j = 0, \ldots, k-1)$. Ainsi

$$v_i = \sum_{j=0}^{i} \varkappa_{ij} (x_j - \mu) \quad (i = 0, \ldots, k-1). \tag{6.238}$$

Par conséquent, $P(\vec{v}_{k-1} \,|\, \vec{u}_{k-1}, \vec{x}_{k-1}, \mu)$ devient une impulsion unité et dans l'expression (6.228) de α_k la variable d'intégration $\vec{v}_{k-1}$ est remplacée en conformité de (6.238). Nous aboutissons ainsi, à un facteur constant près, à l'expression

$$\alpha_k = \int_{\mu=-\infty}^{\infty} \int_{v_k=-\infty}^{\infty} \left[\mu - x_k^* + a_0 v_k + \sum_{\rho=1}^{k} a_\rho \sum_{j=0}^{k-\rho} \varkappa_{k-\rho,\, j} (x_j - \mu) \right]^2 \times$$

$$\times \exp\left\{ -\frac{\mu^2}{2\sigma_\mu^2} - \frac{(v_k - u_k)^2}{2\sigma_g^2} - \right.$$

$$\left. - \frac{\sum_{i=0}^{k-1} \left[\sum_{j=0}^{i} \left(\varkappa_{ij} \left(x_j - \mu - \sum_{\rho=0}^{j} a_\rho u_{j-\rho} \right) \right) \right]^2}{2\sigma_g^2} \right\} d\mu \, dv_k. \tag{6.239}$$

Introduisons les notations :

$$\left.\begin{aligned}
&x_k^{(0)*} = x_k - a_0 u_k - \sum_{\rho=1}^{k} a_\rho \sum_{j=0}^{k-\rho} \varkappa_{k-\rho,\, j}\, x_j;\\
&x_i^{(0)} = \sum_{j=0}^{i} \varkappa_{ij}\left(x_j - \sum_{\rho=0}^{j} a_\rho u_{j-\rho}\right);\ \sum_{j=0}^{i} \varkappa_{ij} = d_i;\\
&F_k = 1 - \sum_{\rho=1}^{k} a_\rho \sum_{j=0}^{k-\rho} \varkappa_{k-\rho,\, j};\ B_{k-1} = \sum_{i=0}^{k-1} \frac{d_i x_i^{(0)}}{\sigma_g^2};\\
&C_{k-1} = \frac{1}{2}\left(\sum_{i=0}^{k-1} \frac{d_i}{\sigma_g^2} + \frac{1}{2\sigma_\mu^2}\right) > 0;\ D_{k-1} = \sum_{i=0}^{k-1} \frac{(x_i^{(0)})^2}{2\sigma_g^2} > 0;\\
&E_k = \frac{B_{k-1}F_k}{2C_{k-1}} - x_k^{(0)*}.
\end{aligned}\right\} \qquad (6.240)$$

Après transformations l'expression (6.239) devient, à un facteur constant près :

$$\alpha_k = \alpha_k(\vec{u}_k, \vec{x}_{k-1}, x_k^*) \equiv \int_{\mu=-\infty}^{\infty} \exp\left\{-\frac{\mu^2}{2\sigma_\mu^2} - \frac{\sum_{i=0}^{k-1}[x_i^{(0)} - d_i\mu]^2}{2\sigma_g^2}\right\} \times$$

$$\times\left[\int_{z_k=-\infty}^{\infty} [F_k\mu - x_k^{(0)*} + a_0 z_k^2] \exp\left\{-\frac{z_k^2}{2\sigma_g^2}\right\} dz_k\right] d\mu, \qquad (6.241)$$

où $z_k = v_k - u_k$. L'intégrale I_k entre crochets est égale, à un facteur constant près, à l'expression

$$I_k \equiv 1 + \frac{(x_k^{(0)*} - F_k\mu)^2}{a_0^2\sigma_g^2}. \qquad (6.242)$$

C'est pourquoi

$$\alpha_k \equiv \int_{\mu=-\infty}^{\infty} \left[1 + \frac{(x_k^{(0)*} - F_k\mu)^2}{a_0^2\sigma_g^2}\right] \times$$

$$\times \exp\left\{-\frac{\mu^2}{2\sigma_\mu^2} - \frac{\sum_{i=0}^{k-1}[x_i^{(0)} - d_i\mu]^2}{2\sigma_g^2}\right\} d\mu \qquad (6.243)$$

ou, après transformations,

$$\alpha_k \equiv \exp\{-D_{k-1}\}\left[\int_{-\infty}^{\infty} \exp\{-C_{k-1}\mu^2 + B_{k-1}\mu\}\, d\mu + \right.$$

$$\left. + \int_{-\infty}^{\infty} \frac{(x_k^{(0)*} - F_k\mu)^2}{a_0^2\sigma_g^2} \exp\{-C_{k-1}\mu^2 + B_{k-1}\mu\}\right] d\mu. \qquad (6.244)$$

En définissant les intégrales entre crochets (cf. [5.26]), on aboutit à l'expression

$$\alpha_k \equiv \exp\left\{-D_{k-1}+\frac{B_{k-1}^2}{4C_{k-1}}\right\}\left(1+\frac{F_k^2}{2a_0^2\sigma_g^2 C_{k-1}}+\frac{E_k^2}{a_0^2\sigma_g^2}\right). \tag{6.245}$$

Dans cette formule E_k qui s'écrit

$$E_k = \frac{B_{k-1}F_b}{2C_{k-1}} - x_k^{(0)*} = \sum_{i=0}^{k-1}\left(\frac{d_iF_k}{2\sigma_g^2 C_{k-1}}+a_{k-1}\right)\sum_{j=0}^{i}\varkappa_{ij}\times$$
$$\times\left(x_j-\sum_{\rho=0}^{i}a_\rho u_{j-\rho}\right)-x_k^*+\sum_{j=0}^{k}a_{k-j}u_j \tag{6.246}$$

est la seule à dépendre de u_k.

Compte tenu de (6.245) on en tire que la minimisation de α_n par rapport à u_n se ramène à la minimisation de E_n par rapport à u_n et cette dernière s'obtient si on choisit un u_n^* tel que $E_n = 0$. Par conséquent,

$$u_n^* = \frac{1}{a_0}\left[x_n^*-\sum_{i=0}^{n-1}\left(\frac{d_iF_n}{2\sigma_g^2 C_{n-1}}+a_{n-i}\right)\sum_{j=0}^{i}\varkappa_{ij}\times\right.$$
$$\left.\times\left(x_j-\sum_{\rho=0}^{j}a_\rho u_{j-\rho}\right)-\sum_{j=0}^{n-1}a_{n-j}u_j\right]=\frac{K_n}{a_0} \tag{6.247}$$

où K_n désigne l'expression entre crochets.

Telle est la stratégie optimale au n-ième pas.

En passant de n à $n-1$, il faut intégrer $\gamma_n^* = \alpha_n^*$ par rapport à x_{n-1}, puis ajouter à l'intégrale la fonction α_{n-1}, après quoi on peut trouver la valeur u_{n-1}^* minimisant l'expression obtenue. La formule (6.245) montre que pour $E_0 = 0$ la fonction $\alpha_n = \alpha_n^*$ dépend de u_{n-1}, cette grandeur faisant partie des expressions de D_{n-1} et B_{n-1}. Pourtant, la substitution

$$x_{n-1}-\sum_{j=0}^{n-1}a_{n-1-j}u_j=\lambda_{n-1} \tag{6.248}$$

et l'intégration par rapport à la nouvelle variable λ_{n-1} de $-\infty$ à $+\infty$ montrent que $\int_{-\infty}^{\infty}\alpha_n^*\,dx_{n-1}$ ne dépend pas de u_{n-1}. De ce fait la minimisation de γ_{n-1} par rapport à u_{n-1} se ramène à la minimisation de α_{n-1} par rapport à u_{n-1}. Des raisonnements analogues montrent que, pour tout k, u_k^* s'obtient en minimisant α_k, ce qui se ramène à une sélection de u_k^* telle qu'elle vérifie la condition $E_k=0$. On en tire la

stratégie optimale pour un s-ième pas quelconque :

$$u_s^* = \frac{1}{a_s}\left[x_s^* - \sum_{i=0}^{s-1}\left(\frac{d_i F_s}{2\sigma_g^2 C_{s-1}} + a_{s-i}\right)\times\right.$$

$$\left.\times\sum_{j=0}^{i}\varkappa_{ij}\left(x_j - \sum_{\rho=0}^{i} a_\rho u_{j-\rho}\right) - \sum_{j=0}^{s-1} a_{s-j}u_j\right] = \frac{K_s}{a_s}. \quad (6.249)$$

où K_s désigne l'expression entre crochets.

Le sens physique du résultat obtenu devient plus clair si l'on tient compte du fait que

$$\mu = x_i - w_i = x_i - \sum_{j=0}^{i} a_{i-j}v_j = x_i - \sum_{j=0}^{i}[a_{i-j}(u_j + g_j)] =$$

$$= \left(x_i - \sum_{j=0}^{i} a_{i-j}u_j\right) - \sum_{j=0}^{i} a_{i-j}g_j. \quad (6.250)$$

Le dernier terme de cette expression est aléatoire et sa variance dépend de i. L'expression entre parenthèses est une moyenne. Ainsi, le deuxième terme de la formule de K_s donne l'estimation de la grandeur aléatoire μ. Les différentes mesures de μ entrent dans cette estimation avec des poids différents du fait que les variances diffèrent d'une mesure à l'autre.

Le dernier terme de la formule de K_s est une moyenne des résultats des actions précédentes u_j ($j < s$) qui reste conservée à l'instant $t = s$ grâce à l'existence d'une mémoire. Pour définir u_s^* il faut évidemment tenir compte de ce reste. Ainsi le sens physique de tous les termes de la formule (6.249) est bien clair.

On peut montrer, sous certaines conditions supplémentaires, que pour un s suffisamment grand le terme médian de la formule de K_s se ramène à la grandeur

$$\frac{1}{s}\sum_{i=0}^{s-1}\left(x_i - \sum_{j=0}^{i} a_{i-j}u_j\right), \quad (6.251)$$

c'est-à-dire à la moyenne arithmétique des mesures de μ.

En principe, on pourrait étendre le résultat obtenu au cas d'un système continu. Si dans le cas continu la relation entre $w(t)$ et $v(t)$ est celle, par exemple, d'un élément avec inertie de fonction de transfert

$$K(p) = \frac{1}{1 + pT_0}, \quad (6.252)$$

on a alors

$$w(t) = \int_0^t b(\tau)\, v(t-\tau)\, d\tau, \quad (6.253)$$

avec la fonction de pondération

$$b(t)=\frac{1}{T_0}\exp\left\{-\frac{t}{T_0}\right\}. \tag{6.254}$$

Rendons discrète l'équation (6.253) en posant

$$w(t)\approx\sum_{\rho=0}^{s} b(t_\rho)\,v(t-t_\rho)\,\Delta t=\sum_{\rho=0}^{s} a_\rho v(t-t_\rho)=\sum_{\rho=0}^{s} a_\rho v_{s-\rho}, \tag{6.255}$$

où

$$a_\rho=b(t_\rho)\,\Delta t=\frac{\Delta t}{T_0}\exp\left\{-\frac{\rho\Delta t}{T_0}\right\},\quad a_0=\frac{\Delta t}{T_0}. \tag{6.256}$$

En portant a_ρ dans la formule (6.249) de la stratégie optimale on pourrait, en faisant tendre Δt vers zéro et en posant $s\Delta t = t$, obtenir à la limite une stratégie optimale pour le cas continu, de même qu'il en a été dans l'exemple exposé au § 2 de ce chapitre. Toutefois, et des considérations physiques en font foi, dans ces conditions $a_0 \to 0$ et, par exemple, la valeur de u_n^* de la formule (6.247) devient infiniment grande. Il en est de même pour d'autres u_s. Or, pour un système réel la seule solution ayant un sens physique est celle qui correspond aux valeurs limitées de u_s. C'est pourquoi dans le cas discret on peut introduire encore une condition supplémentaire, par exemple

$$|u_s|\leqslant M, \tag{6.257}$$

où M est une certaine grandeur constante.

Pourtant, la stratégie optimale se modifie dans ces conditions. Ainsi au lieu de (6.247) il faut maintenant écrire

$$u_n^*=\begin{cases}\dfrac{K_n}{a_0} & \text{si } \left|\dfrac{K_n}{a_0}\right|\leqslant M,\\[2mm] M\operatorname{sign}\dfrac{K_n}{a_0} & \text{si } \left|\dfrac{K_n}{a_0}\right|>M.\end{cases} \tag{6.258}$$

L'exploration de ce problème montre que maintenant $\int \gamma_n^* dx_{n-1}$ dépend de u_{n-1} et le système ne peut plus être considéré comme neutre. Les calculs deviennent donc bien plus compliqués.

Le système considéré est-il réductible? La réponse à cette question est affirmative. La définition d'un système réductible, donnée au paragraphe précédent, peut être étendue également aux systèmes avec mémoire. A cet effet, au lieu de la fonction $f(s, z_s, g_s, h_s)$, caractéristique du bloc E (fig. 6.9, b), il faut considérer une fonction plus complexe $f(s, \vec{z}_s, \vec{g}_s, \vec{h}_s)$, où $\vec{z}_s, \vec{g}_s, \vec{h}_s$ sont les vecteurs temporels respectifs. Dans l'exemple concerné, en vertu de (6.222) et (6.224) et compte tenu de (6.250) on a

$$x_i-\sum_{k=0}^{i} a_k u_{i-k}=\mu+\sum_{k=0}^{i} a_k g_{i-k}. \tag{6.259}$$

Par conséquent, la combinaison des grandeurs mesurées qui figure dans le premier membre de cette égalité, constitue l'équivalent de la mesure de la grandeur μ avec l'erreur aléatoire $\sum_{k=0}^{s} a_k g_{s-k}$. Ainsi, la formation des grandeurs x_i^0 utilisées dans l'algorithme de commande est complètement équivalente à la mesure de la valeur de μ dans un

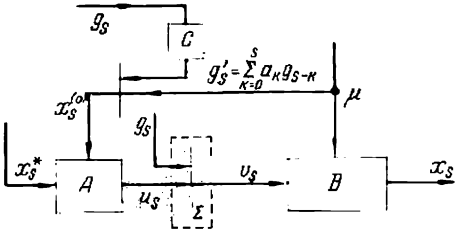


Fig. 6.15.

schéma en boucle ouverte équivalent, représenté sur la figure 6.15. Ce schéma correspond à l'équation (6.259). Le bloc C est caractérisé par l'équation

$$g'_s = \sum_{k=0}^{s} a_k g_{s-k}, \tag{6.260}$$

où g_s est l'entrée et g'_s la sortie du bloc. Ce bloc est identique au bloc B_1 du système.

Le schéma considéré étant réductible, il doit être neutre dans ce sens que la vitesse de l'étude du système B ne doit pas dépendre des valeurs u_s. Quelle est alors la cause du fait qu'en présence de la contrainte (6.257) l'intégrale

$$\int_{-\infty}^{\infty} \gamma_{n-k}^* \, dx_{n-1}$$

dépend quand même de u_{n-k} ce qui complique considérablement les calculs? Il se trouve que cette cause réside dans le fait que pour les systèmes à mémoire l'intégrale caractérise non seulement le risque d'étude, mais encore la partie connue du risque d'action, ou plus précisément, le risque de « postaction ». En effet, la commande u_s appliquée à l'instant $t = s$ exercera son influence sur des x_k pendant les pas suivants ($k > s$), la postaction se produisant également dans le schéma 6.15. Tant que la contrainte (6.257) ne jouait aucun rôle cette postaction pouvait être négligée. En effet, la loi de commande optimale (6.249) comporte un terme (le dernier entre crochets) qui,

précisément, compense toutes les « traces » des actions antérieures conservées à la sortie par la mémoire de l'objet B. Pourtant, si u_s est borné en module, une telle compensation peut devenir impossible. La commande u_s intervient alors également dans les valeurs ultérieures de x_k même dans un système en boucle ouverte (de même que dans un système en boucle fermée équivalent de la fig. 6.14). Il ne faut pas confondre cette influence avec le risque d'étude; dans les systèmes réductibles elle prend la forme d'une composante du risque d'action qu'on peut appeler risque de postaction.

§ 5. Généralisation aux systèmes markoviens

Considérons maintenant la généralisation à une classe plus large des processus aléatoires z_i et x_i^* [6.9, 6.10]. Soient

$$z_s = z_s(s, \bar{\mu}_s) \tag{6.261}$$

et

$$x_i^* = x_i^*(s, \bar{\lambda}_s), \tag{6.262}$$

où, à la différence de l'exposé précédent, $\bar{\mu}_s$ et $\bar{\lambda}_s$ ne sont pas des variables, mais des processus markoviens aléatoires discrets. Les systèmes où $\bar{\mu}_s$ est un processus markovien sont dits systèmes markoviens. Les vecteurs

$$\bar{\mu}_s = (\mu_s^1, \ldots, \mu_s^m) \tag{6.263}$$

et

$$\bar{\lambda}_s = (\lambda_s^1, \ldots, \lambda_s^l) \tag{6.264}$$

sont des ensembles des grandeurs scalaires μ_s^i et λ_s^j ($i = 1, \ldots, m$, $j = 1, \ldots, l$). Dans le cas général, μ_s^i et μ_s^k sont des processus markoviens discrets scalaires associés entre eux. Ceci est vrai également pour λ_s^j et λ_s^r. Pourtant, nous considérerons les vecteurs $\bar{\mu}_s$ et $\bar{\lambda}_s$ ainsi que les bruits h_s^*, h_s, g_s indépendants entre eux.

Soient les caractéristiques probabilistes des processus markoviens $\bar{\mu}_s$ et $\bar{\lambda}_s$. Cela signifie que les densités de probabilité initiales $P_0(\bar{\mu}_0)$ et $P_0(\bar{\lambda}_0)$ pour $t = 0$ sont données de même que les densités de probabilité de transition, c'est-à-dire $P(\bar{\mu}_{i+1} | \bar{\mu}_i)$ et $P(\bar{\lambda}_{i+1} | \bar{\lambda}_i)$. Pour le reste le problème est le même; seulement pour simplifier, bornons-nous au cas où le système B n'a qu'une sortie et une entrée. Supposons que le système soit sans mémoire. La généralisation aux systèmes à mémoire et à entrées et sorties multiples peut se faire de même que précédemment. Pour la généralisation considérée, il faut introduire les matrices $\vec{\bar{\mu}}_s$ et $\vec{\bar{\lambda}}_s$ composées de vecteurs colonnes $\bar{\mu}_s$ et $\bar{\lambda}_s$:

$$\vec{\bar{\mu}}_s = (\bar{\mu}_0, \bar{\mu}_1, \ldots, \bar{\mu}_s); \quad \vec{\bar{\lambda}}_s = (\bar{\lambda}_0, \bar{\lambda}_1, \ldots, \bar{\lambda}_s). \tag{6.265}$$

Le risque élémentaire r_s est donné par la formule très proche des formules précédentes et qui ne s'en distingue que par $\overline{\lambda}_s$ qui vient remplacer $\overline{\lambda}$. Par ailleurs, dans les formules ultérieures, au lieu de $\overline{\mu}$ figure $\overline{\mu}_s$. Ainsi

$$r_s = M\{W_s \mid \vec{y}_s^*, \vec{y}_{s-1}, \vec{u}_{s-1}\} =$$
$$= \int\limits_{\Omega(\overline{\lambda}_s,\, x_s)} W_s[s, x_s, x_s^*(s, \overline{\lambda}_s)]\, P(\overline{\lambda}_s, x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})\, d\Omega. \quad (6.266)$$

Ici $P(\overline{\lambda}_s, x_s \mid \vec{y}_s^*, \vec{u}_{s-1}^*, \vec{y}_{s-1})$ est la densité de probabilité conditionnelle conjointe de $\overline{\lambda}_s$ et x_s pour les vecteurs fixés $\vec{y}_s^*$, $\vec{y}_{s-1}$, $\vec{u}_{s-1}$. D'après le théorème du produit des probabilités

$$P(\overline{\lambda}_s, x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) = P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}, \overline{\lambda}_s)\, P(\overline{\lambda}_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) =$$
$$= P(\overline{\lambda}_s \mid \vec{y}_s^*)\, P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}). \quad (6.267)$$

Cette dernière transformation se vérifie du fait que la densité de probabilité $\overline{\lambda}_s$ avec $\vec{y}_s^*$ fixé ne changera pas si l'on fixe encore $\vec{u}_{s-1}$, $\vec{y}_{s-1}$ (cf. fig. 6.2). Ensuite, la densité de probabilité de x_s avec $\vec{y}_s^*$ fixé ne changera pas si on fixe encore $\overline{\lambda}_s$. Récrivons le deuxième facteur de (6.267) sous une forme développée :

$$P(x_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) =$$
$$= \int\limits_{\Omega(\overline{\mu}_s,\, u_s)} P(x_s \mid \overline{\mu}_s, u_s) \cdot P_s(\overline{\mu}_s) \cdot \Gamma_s(u_s \mid \vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})\, d\Omega. \quad (6.268)$$

$P_s(\overline{\mu}_s)$ est ici la densité de probabilité a priori de $\overline{\mu}_s$ au s-ième pas :

$$P_s(\overline{\mu}_s) = P(\overline{\mu}_s \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s) =$$
$$= \int\limits_{\Omega(\overline{\mu}_{s-1})} P(\overline{\mu}_s \mid \vec{\mu}_{s-1}) \cdot P(\overline{\mu}_{s-1} \mid \vec{y}_{s-1}, \vec{u}_{s-1}, \vec{y}_s^*)\, d\Omega. \quad (6.269)$$

Etant donné que

$$P(\overline{\mu}_{s-1} \mid \vec{y}_{s-1}, \vec{u}_{s-1}, \vec{y}_s^*) = \int\limits_{\Omega(\vec{\overline{\mu}}_{s-2})} P(\vec{\overline{\mu}}_{s-1} \mid \vec{y}_{s-1}, \vec{u}_{s-1}, \vec{y}_s^*)\, d\Omega, \quad (6.270)$$

il faut rechercher la probabilité conditionnelle $P(\vec{\bar{\mu}}_{s-1} \mid \vec{y}_{s-1}, \vec{u}_{s-1}, \vec{y}_s^*)$ pour la matrice $\vec{\bar{\mu}}_{s-1}$. L'égalité

$$P(\vec{\bar{\mu}}_{s-1}, \vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*) = P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{\bar{\mu}}_{s-1}, \vec{y}_s^*)\, P(\vec{\bar{\mu}}_{s-1}) =$$

$$= P(\vec{\bar{\mu}}_{s-1} \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s^*) \cdot P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*) \qquad (6.271)$$

entraîne :

$$P_{s-1}(\vec{\bar{\mu}}_{s-1}) = P(\vec{\bar{\mu}}_{s-1} \mid \vec{u}_{s-1}, \vec{y}_{s-1}, \vec{y}_s^*) =$$

$$= \frac{P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{\bar{\mu}}_{s-1}, \vec{y}_s^*) \cdot P(\vec{\bar{\mu}}_{s-1})}{P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*)}. \qquad (6.272)$$

$P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*)$ est ici la densité de probabilité conditionnelle a priori conjointe des vecteurs $\vec{u}_{s-1}$, $\vec{y}_{s-1}$; $P(\vec{\bar{\mu}}_{s-1})$ la densité de probabilité a priori de la matrice $\vec{\bar{\mu}}_{s-1}$ et $P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{\bar{\mu}}_{s-1}, \vec{y}_s^*)$ la densité de probabilité conditionnelle conjointe de $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$, la matrice $\vec{\bar{\mu}}_{s-1}$ et $\vec{y}_s^*$ (fonction de vraisemblance) étant fixées. Des calculs analogues à ceux effectués au § 1 amènent

$$P_{s-1}(\vec{\bar{\mu}}_{s-1}) = \frac{P_0(\bar{\mu}_0) \cdot \prod_{i=1}^{s-1} P(\bar{\mu}_i \mid \bar{\mu}_{i-1}) \left[\prod_{i=0}^{s-1} P(y_i \mid \bar{\mu}_i, i, u_i)\right] \cdot \left[\prod_{i=0}^{s-1} \Gamma_i\right]}{P(\vec{u}_{s-1}, \vec{y}_{s-1} \mid \vec{y}_s^*)}. \qquad (6.273)$$

En portant (6.273) dans (6.270), puis (6.270) dans (6.269), (6.269) dans (6.268), (6.268) dans (6.267) on obtient le deuxième facteur de (6.267). Considérons maintenant le premier facteur de cette expression, qui est la densité de probabilité a posteriori du vecteur $\bar{\lambda}_s$:

$$P_s(\bar{\lambda}_s) = P(\bar{\lambda}_s \mid \vec{y}_s^*) = \int_{\Omega(\vec{\bar{\lambda}}_{s-1})} P(\vec{\bar{\lambda}}_s \mid \vec{y}_s^*)\, d\Omega. \qquad (6.274)$$

Etant donné que

$$P(\vec{\bar{\lambda}}_s, \vec{y}_s^*) = P(\vec{\bar{\lambda}}_s) \cdot P(\vec{y}_s^* \mid \vec{\bar{\lambda}}_s) = P(\vec{\bar{\lambda}}_s \mid \vec{y}_s^*) \cdot P(\vec{y}_s^*), \qquad (6.275)$$

il vient

$$P_s(\vec{\bar{\lambda}}_s) = P(\vec{\bar{\lambda}}_s) \cdot \frac{P(\vec{y}_s^* \mid \vec{\bar{\lambda}}_s)}{P(\vec{y}_s^*)}. \qquad (6.276)$$

La densité de probabilité a priori $P(\vec{\bar{\lambda}}_s)$ de la matrice $\vec{\bar{\lambda}}_s$ est définie d'après la formule vérifiée pour un processus markovien

$$P(\vec{\bar{\lambda}}_s) = P(\bar{\lambda}_0, \bar{\lambda}_1, \ldots, \bar{\lambda}_s) =$$
$$= P_0(\bar{\lambda}_0) \cdot P(\bar{\lambda}_1 | \bar{\lambda}_0) \cdot P(\bar{\lambda}_2 | \bar{\lambda}_1) \ldots P(\bar{\lambda}_s | \bar{\lambda}_{s-1}) =$$
$$= P_0(\bar{\lambda}_0) \cdot \prod_{i=0}^{s} P(\bar{\lambda}_i | \bar{\lambda}_{i-1}). \qquad (6.277)$$

La chaîne H^* étant sans inertie, la densité de probabilité conditionnelle de $\vec{y}_s^*$ pour $\vec{\bar{\lambda}}_s$ fixée s'obtient de la façon suivante:

$$P(\vec{y}_s^* | \vec{\bar{\lambda}}_s) = P(y_0^* | \bar{\lambda}_0) \cdot P(y_1^* | \bar{\lambda}_1) \ldots P(y_s^* | \bar{\lambda}_s) = \prod_{i=0}^{s} P(y_i^* | \bar{\lambda}_i). \qquad (6.278)$$

(6.276), (6.277) et (6.278) amènent:

$$P_s(\vec{\bar{\lambda}}_s) = \frac{P_0(\bar{\lambda}_0) \cdot \prod_{i=1}^{s} P(\bar{\lambda}_i | \bar{\lambda}_{i-1}) \cdot \prod_{i=0}^{s} P(y_i^* | \bar{\lambda}_i)}{P(\vec{y}_s^*)}. \qquad (6.279)$$

En portant (6.279) dans (6.274) nous aboutissons à la formule définitive de $P_s(\vec{\bar{\lambda}}_s)$.

Après les substitutions indiquées, en portant ensuite (6.267) dans (6.266) on peut obtenir la formule définitive du risque élémentaire conditionnel r_s. Si l'on considère les valeurs de r_s fournies par des épreuves différentes, les vecteurs $\vec{y}_s^*$, $\vec{u}_{s-1}$ et $\vec{y}_{s-1}$ inconnus d'avance peuvent prendre des valeurs différentes et sont aléatoires. Leur densité de probabilité conjointe

$$P(\vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1}) = P(\vec{u}_{s-1}, \vec{y}_{s-1} | \vec{y}_s^*) P(\vec{y}_s^*). \qquad (6.280)$$

Le risque élémentaire R_s, qui est une moyenne de r_s pour le cas d'un très grand nombre d'épreuves, est défini par la formule

$$R_s = M\{r_s\} = \int\limits_{\Omega(\vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})} r_s \cdot P(\vec{y}_s^*, \vec{u}_{s-1}, \vec{y}_{s-1})\, d\Omega =$$
$$= \int\limits_{\Omega(\vec{y}_s, \vec{u}_{s-1}, \vec{y}_{s-1})} r_s P(\vec{u}_{s-1}, \vec{y}_{s-1} | \vec{y}_s^*) P(\vec{y}_s^*)\, d\Omega. \qquad (6.281)$$

En remplaçant r_s par son expression on aboutit à la formule

$$R_s = \int\limits_{\Omega(\vec{\bar{\lambda}}_s, x_s, \vec{\bar{\mu}}_s, \vec{y}_s^*, u_s, \vec{y}_{s-1})} W_s[s, x_s^*(s, \bar{\lambda}_s), x_s] \cdot P_0(\bar{\lambda}_0) \times$$

$$\times \prod_{i=1}^{s} P(\bar{\lambda}_i \mid \bar{\lambda}_{i-1}) \cdot \prod_{i=0}^{s} P(y_i^* \mid \bar{\lambda}_i) \cdot P(x_k \mid k, \bar{\mu}_k, u_k) \cdot P_0(\bar{\mu}_0) \times$$

$$\times \prod_{i=1} P(\bar{\mu}_i \mid \bar{\mu}_{i-1}) \cdot \prod_{i=0}^{s-1} P(y_i \mid \bar{\mu}_i, i, u_i) \cdot \prod_{i=0}^{s} \Gamma_i \, d\Omega. \qquad (6.282)$$

Introduisons les fonctions auxiliaires α_k $(0 \leqslant k \leqslant n)$:

$$\alpha_k = \alpha_k(\vec{y}_k^*, u_k, \vec{u}_{k-1}, \vec{y}_{k-1}) = \int\limits_{\Omega(\vec{\lambda}_k, \vec{\mu}_k, x_k)} W_k[k, x_k^*(k, \bar{\lambda}_k), x_k] \times$$

$$\times P(x_k \mid k, \bar{\mu}_k, u_k) \cdot P_0(\bar{\lambda}_0) \prod_{i=1}^{k} P(\bar{\lambda}_i \mid \bar{\lambda}_{i-1}) \cdot \prod_{i=0}^{k} P(y_i^* \mid \bar{\lambda}_i) \times$$

$$\times P_0(\bar{\mu}_0) \cdot \prod_{i=1}^{k} P(\bar{\mu}_i \mid \bar{\mu}_{i-1} \cdot \prod_{i=0}^{k-1} P(y_i \mid \bar{\mu}_i, i, u_i) \, d\Omega$$

$$(k = 0, 1, \ldots, n). \qquad (6.283)$$

Soit également

$$\beta_k = \prod_{i=0}^{k} \Gamma_i. \qquad (6.284)$$

Dans ce cas la formule du risque R_n, par exemple, associé à l'instant $t = n$, se mettra sous la même forme que (6.38). En reprenant les raisonnements du § 1, on aboutit à une procédure analogue de la définition de la stratégie optimale: soit $\gamma_n = \alpha_n$ et

$$\gamma_{n-k} = \gamma_{n-k}(\vec{y}_{n-k}^*, \vec{u}_{n-k}, \vec{y}_{n-k-1}) =$$

$$= \alpha_{n-k} + \int\limits_{\Omega(y_{n-k}, y_{n-k+1}^*)} \gamma_{n-k+1}^*(u_{n-k+1}^*, \vec{y}_{n-k+1}, \vec{u}_{n-k}, \vec{y}_{n-k}^*) \, d\Omega$$

$$(k = 0, 1, \ldots, n). \qquad (6.285)$$

De plus,

$$\gamma_{n-k}^* = \min_{u_{n-k} \in \Omega(u_{n-k})} \gamma_{n-k} = \gamma_{n-k}(u_{n-k}^*). \qquad (6.286)$$

Evidemment,

$$u_{n-k}^* = u_{n-k}^*(\vec{y}_{n-k}^*, \vec{u}_{n-k-1}, \vec{y}_{n-k-1}). \qquad (6.287)$$

La stratégie optimale est alors donnée par l'expression

$$\Gamma_{n-k}^* = \delta(u_{n-k} - u_{n-k}^*). \qquad (6.288)$$

qui montre que c'est une stratégie pure consistant à choisir $u_{n-k} = u^*_{n-k}$.

Dans le cas particulier, lorsque le processus $\bar{\lambda}_s$ se transforme en variable $\bar{\lambda}$ et $\bar{\mu}_s$ se transforme en variable $\bar{\mu}$, la formule (6.283) de α_k devient plus simple:

$$\alpha_k = \int\limits_{\Omega(\bar{\lambda},\, \bar{\mu},\, x_k)} W_k\,[k,\ x_k^*(k,\ \bar{\lambda}),\ x_k]\cdot P_0(\lambda)\times$$

$$\times \prod_{i=0}^{k} P(y_i^* \mid \bar{\lambda})\, P_0(\bar{\mu}) \prod_{i=0}^{k-1} P(y_i \mid \bar{\mu},\ i,\ u_i)\, d\Omega. \qquad (6.289)$$

Cette formule coïncide avec (6.36).

Si la fonction $\bar{\mu}$ ne se ramène pas à une grandeur constante et constitue un processus markovien, on a intérêt de résoudre le problème de la commande duale d'un processus stable. Dans le cas contraire,

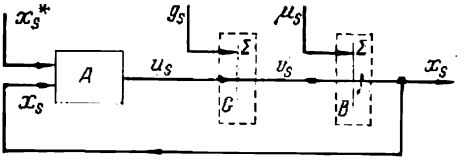


Fig. 6.16.

si la grandeur μ est constante tout au cours du processus et si le nombre de mesures est infiniment grand, vers la fin du processus on peut connaître la valeur de μ avec une erreur aussi petite que l'on veut et alors l'information sur le système devient complète.

Soit $n \to \infty$. Dans un processus stable on peut imposer le choix d'une stratégie telle qu'elle minimise le risque moyen relatif à un pas, c'est-à-dire la grandeur

$$\rho = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} R_i. \qquad (6.290)$$

S'il existe une limite

$$R_\infty = \lim_{\substack{i\to\infty \\ n\to\infty \\ n-i\to\infty}} R_i, \qquad (6.291)$$

alors $\rho = R_\infty$.

Pour illustrer la théorie exposée ci-dessus, considérons à titre d'exemple un système dont le schéma fonctionnel est donné par la figure 6.16. Soient le système B et la chaîne G d'équations

$$x_s = \mu_s + v_s \qquad (6.292)$$

et

$$v_s = u_s + g_s \tag{6.293}$$

respectivement. μ_s est ici un processus markovien gaussien discret de densité de probabilité initiale

$$P_0(\mu_0) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left\{-\frac{\mu_0^2}{2\sigma_0^2}\right\} \tag{6.294}$$

et de densité de probabilité de transition

$$P(\mu_k \mid \mu_{k-1}) = \frac{1}{\sigma_1 \sqrt{2\pi}} = \exp\left\{-\frac{(\mu_k - \mu_{k-1})^2}{2\sigma_1^2}\right\}. \tag{6.295}$$

Les grandeurs g_s sont indépendantes et $P(g_s)$ s'écrit:

$$P(g_s) = q(g_s) = \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{g_s^2}{2\sigma_g^2}\right\}. \tag{6.296}$$

Soit, ensuite, $x_s^* = x^* = \text{const}$ et

$$W_k = (x^* - x_k)^2. \tag{6.297}$$

Puisque $h_s = 0$, $y_s = x_s$. Dans cet exemple

$$P(y_i \mid \mu_i, i, u_i) = P(x_i \mid \mu_i, u_i) =$$
$$= q(x_i - u_i - \mu_i) = \frac{1}{\sigma_g \sqrt{2\pi}} \exp\left\{-\frac{(x_i - u_i - \mu_i)^2}{2\sigma_g^2}\right\}. \tag{6.298}$$

α_k pour l'exemple considéré est fournie par l'expression générale (6.283). Il faut tenir compte qu'ici $h_s^* = 0$; donc $P_0(\bar{\lambda}_0)$, $P(\bar{\lambda}_i \mid \bar{\lambda}_{i-1})$, $P(y_i^* \mid \bar{\lambda}_i)$ dégénèrent en fonctions δ et (6.283) peut se mettre sous la forme

$$\alpha_k = \alpha_k(x_k^*, \vec{u}_k, \vec{y}_{k-1}) = \int\limits_{\Omega(\vec{\bar{\lambda}}_k, \vec{\bar{\mu}}_k, x_k)} W_k(k, x_k^*, x_k) \times$$
$$\times P(x_k \mid k, \bar{\mu}_k, u_k) P_0(\bar{\mu}_0) \cdot \prod_{i=1}^{k} P(\bar{\mu}_i \mid \bar{\mu}_{i-1}) \times$$
$$\times \prod_{i=0}^{k-1} P(y_i \mid \bar{\mu}_i, i, u_i)\, d\Omega. \tag{6.299}$$

En portant ici les expressions correspondantes données par (6.294), (6.295) et (6.298), on a

$$\alpha_k = C \int\limits_{x_k=-\infty}^{\infty} \int\limits_{\mu_0=-\infty}^{\infty} \dots \int\limits_{\mu_k=-\infty}^{\infty} (x^* - x_k)^2 \times$$
$$\times \exp\left\{-\frac{\mu_0^2}{2\sigma_0^2} - \frac{(\mu_1 - \mu_0)^2 + (\mu_2 - \mu_1)^2 + \dots + (\mu_k - \mu_{k-1})^2}{2\sigma_1^2}\right\} \times$$
$$\times \exp\left\{-\sum_{i=0}^{k} \frac{(x_i - u_i - \mu_i)^2}{2\sigma_g^2}\right\} d\mu_0 \cdot d\mu_1 \dots d\mu_k \cdot dx_k, \tag{6.300}$$

où C est une constante. Après le calcul l'intégrale (6.300) devient [6.11] :

$$\frac{\alpha_k}{C'} = \exp\{-\bar{A}^* K \bar{A}\} \left\{ \left[A_k^* - \sum_{i=0}^{k-1} e_{ik} A_i \right]^2 + d_k \right\}. \qquad (6.301)$$

Ici $C' = \text{const}$ et les coefficients sont fournis par les égalités

$$\left.\begin{aligned} &\frac{x_i - u_i}{\sigma_g} = A_i, \quad \frac{\mu_i}{\sigma_g} = \nu_i, \qquad (i = 0, 1, \ldots, k), \\ &\frac{x^* - u_k}{\sigma_g} = A_k^*, \quad \frac{\sigma_0}{\sigma_g} = \vartheta_0, \quad \frac{\sigma_1}{\sigma_g} = \vartheta_1 \end{aligned}\right\} \qquad (6.302)$$

et

$$\left.\begin{aligned} &a_1 = \frac{\vartheta_1^2}{\vartheta_0^2} + \vartheta_1^2 + 1, \\ &a = 2 + \vartheta_1^2, \quad b = 1 + \vartheta_1^2. \end{aligned}\right\} \qquad (6.303)$$

Introduisons de plus les notations des vecteurs A et ν de dimension $(k+1)$:

$$A = (A_0, \ldots, A_k); \quad \nu = (\nu_0, \ldots, \nu_k), \qquad (6.304)$$

et désignons par (A, A) le produit scalaire du vecteur par lui-même, par (A, ν) le produit scalaire des vecteurs A et ν et, enfin, par A^* le vecteur transposé de A (le vecteur A est un vecteur colonne et A^* un vecteur ligne). Soit, ensuite,

$$\bar{A} = (A_0, \ldots, A_{k-1}) \qquad (6.305)$$

le vecteur de dimension k. L'expression $\{\bar{A}^* K \bar{A}\}$ est un produit scalaire de $\bar{A}^*$ et de $K\bar{A}$, où K est une certaine matrice aux coefficients k_{ij} ne contenant pas A_k^*. Nous n'écrirons plus ces coefficients, car par la suite nous n'en aurons pas besoin. Les coefficients e_{ik} et d_k sont définis par les formules

$$e_{ik} = \frac{C_{kk}}{1 - C_{kk}^2} C_{ik}, \quad d_k = \frac{1}{1 - C_{kk}^2}, \qquad (6.306)$$

où C_{ik} sont les coefficients de la matrice triangulaire C (de C_{00} à C_{kk}) déterminés par les expressions

$$C_{jk} = \frac{\Delta_j \vartheta_1}{\sqrt{\Delta_k (b\Delta_k - \Delta_{k-1})}} = \frac{\Delta_j}{\Delta_k} C_{kk} \qquad (0 \leqslant j \leqslant k). \qquad (6.307)$$

Dans ces conditions nous supposons $\Delta_0 = 1$ et

$$C_{kk} = \vartheta_1 \sqrt{\frac{\Delta_k}{b\Delta_k - \Delta_{k-1}}}. \qquad (6.308)$$

alors que Δ_i $(i = 1, \ldots, k)$ est le déterminant à i lignes et aux colonnes

$$\Delta_i = \begin{vmatrix} a & -1 & 0 & & & 0 \\ -1 & a & -1 & & & \\ 0 & -1 & a & & & \\ & & & \ddots & & \\ & & & a & -1 & 0 \\ & & & -1 & a & -1 \\ 0 & & & 0 & -1 & a \end{vmatrix}. \tag{6.309}$$

En portant (6.307) dans la première des formules (6.306), on tombe sur

$$e_{ik} = \frac{1}{\frac{(b\Delta_k - \Delta_{k-1})}{\vartheta_1^2 \Delta_k} - 1} \cdot \frac{\Delta_i}{\Delta_k}. \tag{6.310}$$

Posons

$$Q_i = \frac{\Delta_i}{\Delta_{i-1}}, \quad q = \frac{b\Delta_k - \Delta_{k-1}}{\Delta_k}. \tag{6.311}$$

Il est évident que $Q_1 = \frac{\Delta_1}{\Delta_0} = \Delta_1 = a_1 > 1$.

Si l'on tient compte de (6.311), la formule (6.310) peut s'écrire

$$e_{ik} = \frac{1}{\frac{q}{\vartheta_1^2} - 1} \cdot \frac{\Delta_i}{\Delta_k}. \tag{6.312}$$

On montre facilement que (6.309) entraîne

$$\Delta_i = a\Delta_{i-1} - \Delta_{i-2} \qquad (i = 2, \ldots, k). \tag{6.313}$$

On en tire

$$Q_i = a - \frac{1}{Q_{i-1}} \qquad (i = 2, \ldots, k). \tag{6.314}$$

Etant donné que $Q_1 > 1$ et $a = 2 + \vartheta_1^2 > 2$, (6.314) amène

$$Q_i > 1 \qquad (i = 1, \ldots, k). \tag{6.315}$$

La grandeur

$$q = b - \frac{1}{Q_k} > 0, \tag{6.316}$$

puisque $b = 1 + \vartheta_1^2 > 1$; mais q peut être soit plus grand soit plus petit que l'unité.

Si l'on porte dans (6.301) les expressions de A_i et A_k^* conformément à (6.302), on obtient pour $k = n$, par exemple,

$$\alpha_n = \left\{\left[\frac{x^* - u_n}{\sigma_g} - \sum_{i=0}^{n-1} e_{in}\left(\frac{x_i - u_i}{\sigma_g}\right)\right]^2 + d_n\right\} \times$$
$$\times \exp\left\{-\sum_{j=0}^{n-1} \frac{k_{ij}}{2} \cdot \frac{x_i - u_i}{\sigma_g} \cdot \frac{x_j - u_j}{\sigma_g}\right\}. \qquad (6.317)$$

La grandeur u_n ne fait partie que de l'expression entre crochets. Pour minimiser α_n par rapport à u_n il faut annuler l'expressiou considérée. Cette condition entraîne

$$u_n^* = x^* - \sum_{i=0}^{n-1} e_{in}(x_i - u_i). \qquad (6.318)$$

D'après la procédure générale de la définition d'une stratégie optimale, la valeur de u_{n-1}^* s'obtient si l'on calcule d'abord γ_{n-1}:

$$\gamma_{n-1} = \alpha_{n-1} + \int_{-\infty}^{\infty} \alpha_n^* \, dx_{n-1} = \alpha_{n-1} + \sigma_g \int_{-\infty}^{\infty} \alpha_n^* \, dA_{n-1}. \qquad (6.319)$$

Dans la dernière intégrale l'expression sous le signe somme s'écrit

$$\alpha_n^* = \min_{u_n} \alpha_n = d_n \exp\left\{-\sum_{i,j=0}^{n-1} \frac{k_{ij}}{2} A_i A_j\right\}. \qquad (6.320)$$

Si l'on porte cette expression dans la formule (6.319), on établit après intégration que le deuxième terme de (6.319) ne dépend pas de A_{n-1} et, par conséquent, ne dépend non plus de u_{n-1}. C'est pourquoi la valeur optimale de u_{n-1}^* peut s'obtenir en minimisant par rapport à u_{n-1} seulement la composante α_{n-1} de l'expression de γ_{n-1}. Dans la formule de α_{n-1}, u_{n-1} n'intervient que dans le terme A_{n-1} de l'expression entre crochets [cf. (6.301)]. En égalant l'expression entre crochets à zéro, on a

$$u_{n-1}^* = x^* - \sum_{i=0}^{n-2} e_{i,\,n-1}(x_i - u_i). \qquad (6.321)$$

Des raisonnements analogues appliqués à $k = n - 2, n - 3, \ldots$ conduisent à la formule générale de la stratégie optimale:

$$u_k^* = x^* - \sum_{i=0}^{k-1} e_{ik}(x_i - u_i). \qquad (6.322)$$

Elucidons le sens physique de cette formule. D'après la figure 6.16 l'espérance mathématique de $(x^* - x_k)^2$ est minimale si l'on choisit

$u_k = x^* - \mu_k$. Le deuxième terme de la formule (6.322) est l'estimation de μ_k basée sur la mesure des différences $(x_i - u_i)$ avec $i < k$. Les coefficients e_{ik} faisant partie de cette estimation dépendent de $(k - i)$. Les valeurs de ces coefficients sont données par la formule (6.310) ou (6.312).

Considérons la relation

$$\frac{e_{ik}}{e_{jk}} = \frac{\Delta_i}{\Delta_j} < 1 \qquad (0 \leqslant i < j \leqslant k). \tag{6.323}$$

Le sens physique de la propriété d'une stratégie optimale traduite par la relation (6.323) consiste dans le fait qu'une information d'origine plus vieille joue un rôle moins important, elle « vieillit ». Ainsdans un organe de commande non seuleument s'accumule une nouvelle information, mais encore s'efface l'information vieillie.

Posons

$$i = k - \nu \tag{6.324}$$

et

$$e_{ik} = \frac{1}{\frac{q}{\vartheta_1^2} - 1} \cdot \frac{\Delta_{k-\nu}}{\Delta_k} = f_{k\nu} \qquad (0 \leqslant i \leqslant k,\ \nu = 0,\ 1,\ \ldots,\ k). \tag{6.325}$$

Recherchons les valeurs des coefficients pour un régime stable

$$\lim_{k \to \infty} f_{k\nu} = f_\nu. \tag{6.326}$$

Examinons au préalable la formule (6.314). Retenons que $Q_1 = a_1 > \frac{a}{2}$. Soit, en général, $Q_{i-1} > \frac{a}{2}$. On a alors

$$Q_i = a - \frac{1}{Q_{i-1}} > a - \frac{2}{a} = \frac{a}{2}\left(2 - \frac{4}{a^2}\right) > \frac{a}{2}. \tag{6.327}$$

Ainsi, tous Q_i vérifient la condition $Q_i > \frac{a}{2}$. D'après (6.314) et (6.327), la relation entre Q_i et Q_{i-1} interprétée géométriquement est une hyperbole d'équation $\left(-\frac{1}{Q_{i-1}}\right)$, déplacée en haut de la grandeur a (courbe *1* de la figure 6.17). La bissectrice du premier quadrant, c'est-à-dire la droite $Q_i = Q_{i-1}$ désignée par le chiffre 2, coupe la courbe *1* en deux points M_1 et M_2. Les abscisses (ou ordonnées) x_1 et x_2 de ces points sont les racines de l'équation quadratique

$$x = a - \frac{1}{x}, \tag{6.328}$$

i.e. les grandeurs

$$x_{1,2} = \frac{a}{2} \pm \sqrt{\frac{a^2}{4} - 1} \tag{6.329}$$

Il s'ensuit que l'ordonnée du point M_1 est inférieure à $\frac{a}{2}$, fait matérialisé par la figure 6.17. Constatons que la construction de ce dessin se vérifie pour tout $i = 1, 2, \ldots$

Imposons-nous une certaine valeur définie de $Q_{i-1} > \frac{a}{2}$.

Sur la figure 6.17, c'est Q'_{i-1} ou Q''_{i-1}. Les valeurs respectives Q_i, Q_{i+1}, Q_{i+2}, etc. peuvent alors être établies graphiquement, à l'aide de l'« échelle » représentée sur la figure 6.17. La valeur Q_{i+1} associée à Q'_i s'obtient si l'on dresse en pointillé une ordonnée depuis le point Q'_i jusqu'à son intersection avec la courbe *1*. L'ordonnée du point d'intersection obtenu, que nous allons appeler premier point de l'« échelle », est égale à Q_{i+1}. Si nous voulons connaître Q_{i+2}, la même valeur doit être portée en abscisses. Pourtant, le report de la valeur Q_{i+1} sur l'axe des abscisses peut être remplacé par le tracé d'une droite horizontale (munie d'une flèche) allant du premier point de l'«échelle » à la bissectrice, c'est-à-dire à la ligne *2*. Le point d'intersection obtenu est le deuxième point de l'« échelle ». On voit aisément que la valeur Q_{i+2} s'obtient si l'on dresse une verticale depuis le deuxième point de l'« échelle » jusqu'à son intersection avec la courbe *1*. Une procédure analogue permet de trouver Q_{i+3}, Q_{i+4}, etc. Comme le montre la figure 6.17, pour $i \to \infty$ la valeur Q_i converge vers Q_∞ égale à la valeur maximale de x. Ainsi, (6.329) conduit à

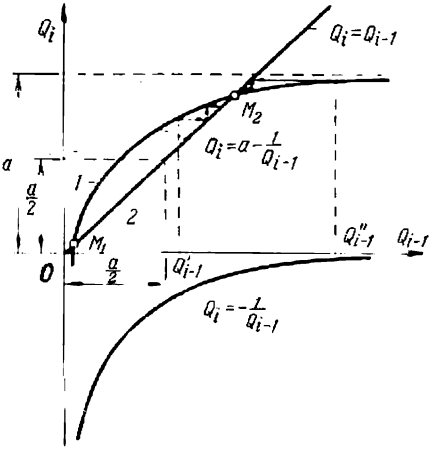


Fig. 6.17.

$$Q_\infty = \lim_{i \to \infty} Q_i = \frac{a}{2} + \sqrt{\frac{a^2}{4} - 1} > \frac{a}{2} > 1. \tag{6.330}$$

Ensuite, pour $k \to \infty$, la grandeur

$$q = b - \frac{1}{Q_k} \to b - \frac{1}{Q_\infty} = q_\infty. \tag{6.331}$$

Par conséquent,

$$f_0 = \frac{1}{\frac{q_\infty}{\vartheta_1^2} - 1} = \frac{1}{\frac{1}{\vartheta_1^2}\left(b - \frac{1}{Q_\infty}\right) - 1}. \tag{6.332}$$

Ensuite

$$\frac{f_{\nu+1}}{f_\nu} = \lim_{k\to\infty} \frac{\Delta_{k-\nu-1}}{\Delta_{k-\nu}} = \lim_{k\to\infty} \frac{1}{Q_{k-\nu}} = \frac{1}{Q_\infty} < 1. \qquad (6.333)$$

Si conformément à (6.325) on substitue $f_{k\nu}$ à e_{ik} pour passer à la limite avec $k \to \infty$, l'expression (6.322) est remplacée par une suite infinie dans laquelle u_0 est une valeur courante et $x_{-\nu}$ et $u_{-\nu}$ sont des valeurs mesurées ν pas auparavant :

$$u_0 = x^* - \sum_{\nu=1}^{\infty} f_\nu (x_{-\nu} - u_{-\nu}). \qquad (6.334)$$

D'après (6.333) les poids f_ν, avec lesquels les différences $x_{-\nu} - u_{-\nu}$ font partie de la somme (6.334), diminuent à mesure que ν croît suivant une progression géométrique.

La formule (6.334) peut être mise sous une forme plus simple. (6.334) et (6.333) entraînent

$$u_{-1} = x^* - \sum_{\nu=2}^{\infty} f_{\nu-1} (x_{-\nu} - u_{-\nu}) = x^* - Q_\infty \sum_{\nu=2}^{\infty} f_\nu (x_{-\nu} - u_{-\nu}). \qquad (6.335)$$

Mais, d'après (6.334), on peut écrire :

$$\sum_{\nu=2}^{\infty} f_\nu (x_{-\nu} - u_{-\nu}) = x^* - u_0 - f_1 (x_{-1} - u_{-1}). \qquad (6.336)$$

En portant (6.336) dans (6.335) on obtient l'égalité

$$Q_\infty u_0 = (Q_\infty - 1)\, x^* + u_{-1} - f_1 (x_{-1} - u_{-1}). \qquad (6.337)$$

On tire de (6.333) que

$$f_1 = \frac{f_0}{Q_\infty}.$$

On tire de (6.337)

$$u_0 = \frac{1}{Q_\infty} \left[(Q_\infty - 1)\, x^* + \left(1 + \frac{f_0}{Q_\infty}\right) u_{-1} - \frac{f_0}{Q_\infty} x_{-1} \right]. \qquad (6.338)$$

La figure 6.18 représente un schéma fonctionnel qui correspond à la formule (6.338). Dans ce schéma, τ est une chaîne de retard d'un pas. De ce fait, lorsque à l'entrée de la chaîne on a u_0, à la sortie on a u_{-1}. Les autres circuits du schéma assument l'amplification et l'un d'entre eux la sommation. Le schéma d'un organe de commande optimal A, prévu seulement pour un régime stable, est muni d'une réaction positive interne réalisée par la chaîne d'amplification $1 + f_0/Q_\infty$ et la chaîne de retard.

Soit, par exemple, $\vartheta_1 = 1$. Alors $b = 1 + \vartheta_1^2 = 2$; $a = 2 + \vartheta_1^2 = 3$. La valeur de Q_∞ calculée d'après l'égalité (6.330) est

égale à 2,615. Il vient

$$f_0 = \frac{1}{\frac{1}{1}\left(2\frac{1}{2,615}\right)-1} = 1,62,$$

$$f_1 = \frac{f_0}{Q_\infty} = \frac{1,62}{2,615} = 0,618,$$

$$f_2 = \frac{f_1}{Q_\infty} = \frac{0,618}{2,615} = 0,235, \text{ etc.}$$

La série (6.334) se met donc sous la forme

$$u_0 = x^* - 0,618\,(x_{-1} - u_{-1}) - 0,235\,(x_{-2} - u_{-2}) - \ldots$$

Passons à la mesure continue.

Soit le temps $T = n\Delta t$ fixé et découpé en n intervalles égaux d'une durée Δt. Supposons d'abord que les mesures soient prises aux

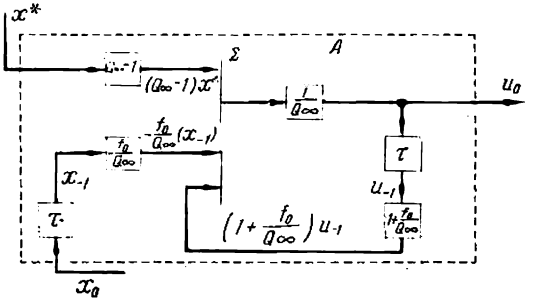


Fig. 6.18.

extrémités des intervalles, alors que Δt intervient dans la valeur de σ_g^2 de la façon suivante

$$\sigma_g^2 = \frac{S_0}{\Delta t}\,. \tag{6.339}$$

La suite des variables g_s donne à la limite pour $\Delta t \to 0$ un bruit blanc de densité spectrale S_0. Soit, ensuite, la variance σ_1^2 de la densité de probabilité de transition dépendant de Δt suivant la loi

$$\sigma_1^2 = \sigma_2^2\,\Delta t, \tag{6.340}$$

avec $\sigma_2^2 = \text{const}$. La formule (6.295) se met alors sous la forme d'une formule de la densité de probabilité d'un processus markovien continu normal

$$P(\mu_0 + \Delta\mu,\ t_0 + \Delta t \mid \mu_0,\ t_0) = \frac{1}{\sigma_2\sqrt{2\pi\Delta t}}\exp\left\{-\frac{(\Delta\mu)^2}{2\sigma_2^2\Delta t}\right\}, \tag{6.341}$$

μ_0 étant la valeur de μ pour $t = t_0$ et $\Delta\mu$ l'accroissement de μ en un temps Δt. Dans ce cas, lorsque $\Delta t \to 0$,

$$\vartheta_1^2 = \frac{\sigma_2^2(\Delta t)^2}{S_0} \to 0, \quad \vartheta_0^2 = \frac{\sigma_0^2\Delta t}{S_0} \to 0. \tag{6.342}$$

Comme pour de petites valeurs de Δt et ϑ_1 la grandeur $Q_\infty \cong \cong 1+\vartheta_1$ on obtient d'après (6.332)

$$f_0 \cong \frac{1}{\frac{1}{\vartheta_1^2}\left(1-\frac{1}{1+\vartheta_1}\right)-1} \cong \vartheta_1 = \frac{\sigma_2}{\sqrt{S_0}}\Delta t = L\Delta t, \quad (6.343)$$

avec

$$L = \frac{\sigma_2}{\sqrt{S_0}}. \quad (6.344)$$

L'expression (6.334) devient

$$u_0 = x^* - \sum_{\nu=1}^{\infty} \frac{L\Delta t}{(1+L\Delta t)^\nu}(x_{-\nu}-u_{-\nu}). \quad (6.345)$$

En désignant les valeurs courantes par $u(t)$ et $x(t)$, on a à la limite pour $\Delta t \to 0$:

$$u(t) = x^* - L\int_{\tau=0}^{\infty} \exp\{-L\tau\}\,[x(t-\tau)-u(t-\tau)]\,d\tau. \quad (6.346)$$

Ainsi pour obtenir une loi de commande optimale il convient d'appliquer la différence $x - u$ à l'entrée d'une chaîne avec inertie de constante de temps L^{-1}. La grandeur de sortie de cette chaîne est définie par le deuxième terme de l'expression (6.346).

§ 6. Schémas fonctionnels des organes de commande optimaux

La comparaison des systèmes automatiques en boucle ouverte et en boucle fermée à objets commandés avec inertie présente un intérêt certain. Pour faire cette comparaison, examinons un cas particulier simple, celui du système en boucle fermée de la figure 6.19.

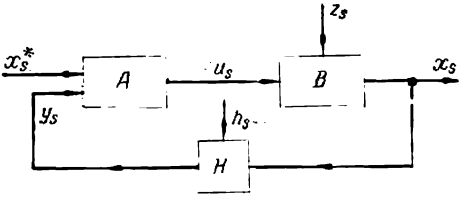


Fig. 6.19.

oit la perturbation z_s assujettie à s et au vecteur des paramètres μ:

$$z_s = z_s(s, \vec{\mu}), \quad (6.347)$$

alors que l'équation de l'objet discret avec inertie peut s'écrire

$$x_s = F_0(\vec{z_s}, \vec{u_s}) = F(\vec{\mu}, \vec{u_s}). \quad (6.348)$$

Nous considérons, en outre, que les conditions initiales sont connues et qu'elles font partie de l'expression (6.348) sous la forme de paramètres. Si ces conditions ne sont pas connues, on peut les associer au vecteur $\bar{\mu}$. La fonction de pertes élémentaire adoptée (en considérant que x_s^* est un paramètre connu) est de la forme

$$W_s = W_s[s, x_s, x_s^*] = W_s[s, F(s, \bar{\mu}, \vec{u}_s)\, x_s^*] = W_s[s, \bar{\mu}, \vec{u}_s]. \quad (6.349)$$

La caractéristique de la chaîne sans inertie H faisant partie de la chaîne de réaction est définie par l'expression

$$y_s = y_s(h_s, x_s). \quad (6.350)$$

Soient $P(\bar{\mu})$ les densités de probabilité a priori du vecteur $\bar{\mu}$ et $P(h_s)$ des variables aléatoires indépendantes h_s ($s = 0, 1, 2, \ldots, n$). Supposons que $\bar{\mu}$ et h_s sont indépendants. La tâche consiste à rechercher la stratégie optimale de l'organe A

$$P_s(u_s) = \Gamma_s(u_s \mid \vec{u}_{s-1}, \vec{y}_{s-1}). \quad (6.351)$$

Le risque élémentaire R_s s'exprime par la formule

$$R_s = M\{W_s\} = \int\limits_{\Omega(\bar{\mu}, \vec{u}_s, \vec{y}_{s-1})} W_s(s, \vec{\mu}, \vec{u}_s) \cdot P(\vec{\mu}) \times \\ \times \prod_{i=0}^{s-1} P(y_i \mid i, \bar{\mu}, \vec{u}_i) \cdot \prod_{i=0}^{s} \Gamma_i \, d\Omega. \quad (6.352)$$

Puisque avec les u_i données l'information sur y_i n'augmente pas si l'on donne en outre $\vec{y}_{i-1}$, l'influence de cette dernière peut ne pas être prise en considération. La fonction α_k s'écrit alors

$$\alpha_k(\vec{u}_k, \vec{y}_{k-1}) = \int\limits_{\Omega(\bar{\mu})} W_k[k, \bar{\mu}, \vec{u}_k] \cdot P(\bar{\mu}) \times \\ \times \prod_{i=0}^{k-1} P(y_i \mid i, \bar{\mu}, \vec{u}_i)\, d\Omega. \quad (6.353)$$

Si l'on pose $\gamma_n = \alpha_n$ et

$$\gamma_{n-k}(\vec{u}_{n-k}, \vec{y}_{n-k-1}) = \alpha_{n-k} + \int\limits_{\Omega(\vec{y}_{n-k})} \gamma_{n-k+1}^* \, d\Omega, \quad (6.354)$$

la définition de la valeur optimale u_s^* se ramène à la minimisation de γ_s par rapport à u_s, ce que nous avons déjà montré à plus d'une reprise.

Peut-on remplacer le système en boucle fermée de la figure 6.19 par un système quelconque en boucle ouverte équivalent sans chaînes de réaction mais muni d'un circuit de mesure de la perturbation z_s?

Dans un schéma équivalent les valeurs de u_s doivent être les mêmes que dans un schéma réel; c'est pourquoi dans les deux cas le

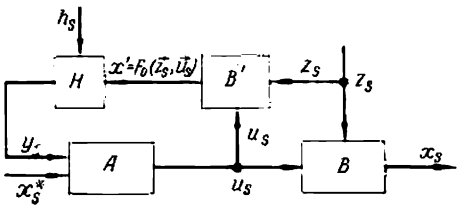


Fig. 6.20.

comportement du système B doit être parfaitement identique. Le schéma de la figure 6.19 se transforme aisément en utilisant la formule déduite de (6.350) et (6.348):

$$y_s = y_s[h_s, x_s] = y_s[h_s, F_0(\vec{z}_s, \vec{u}_s)]. \tag{6.355}$$

En effet, si l'opérateur F_0 de l'objet B est donné, on peut construire son modèle artificiel B' ayant le même opérateur et appliquer z_s

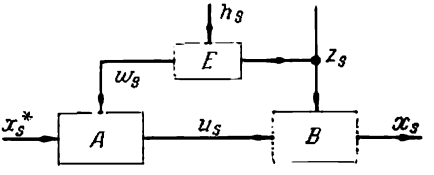


Fig. 6.21.

et u_s à l'entrée du modèle (fig. 6.20). Supposons que les conditions initiales du modèle B' et du système B coïncident. On obtient alors à la sortie du modèle la grandeur x'_s identique à la grandeur de sortie réelle du système.

Dans le cas général, le schéma 6.20 n'est pourtant pas en boucle ouverte, car il comporte une réaction locale qui va de la sortie de l'organe A vers son entrée à travers les chaînes B' et H. Le problème consiste ainsi à établir si l'on peut concevoir un système en boucle ouverte sans aucune réaction extérieure, comme celui représenté sur la figure 6.21, qui serait équivalent au système optimal donné par la figure 6.19? Le bloc E du schéma de la figure 6.21 peut être également un système dynamique, c'est-à-dire sa sortie w_s peut dépendre

non seulement des valeurs courantes de h_s et z_s mais encore de la « préhistoire » des grandeurs d'entrée. Si le système E est sans inertie par rapport à h_s, alors

$$w_s = w_s(h_s, \vec{z}_s) = w_s(h_s, s, \bar{\mu}). \tag{6.356}$$

Le schéma de la figure 6.21 présente également un intérêt particulier. Cherchons pour ce schéma la stratégie optimale de l'organe de commande A sous la forme

$$P_s(u_s) = \Gamma_s(u_s \mid \vec{u}_{s-1}, \vec{w}_{s-1}). \tag{6.357}$$

Explicitons le sens des vecteurs qui se trouvent à droite de la verticale de l'expression (6.357). La valeur optimale lors de la répartition de la grandeur u_s doit dépendre de toutes les valeurs $w_0, w_1, \ldots$ $\ldots, w_{s-1}$, c'est-à-dire du vecteur $\vec{w}_{s-1}$. En effet, ces valeurs permettent de calculer $\vec{z}_s$ et avec elle le vecteur inconnu des paramètres $\bar{\mu}$. Le système E peut être considéré comme un circuit apériodique de mesure de z_s donnant lieu également au bruit h_s. En général, plus les valeurs de w_i ($i = 0, 1, \ldots, s-1$) observées dans le passé sont nombreuses, plus la définition de la valeur du vecteur inconnu des paramètres $\bar{\mu}$ est précise.

Tout autre est le sens du vecteur $\vec{u}_{s-1}$ de l'expression (6.357). Pour assurer une commande correcte du système B il faut en tenir compte lors du calcul de u_s. Etant donné que cet objet est inertiel, à l'instant $t = s$ se manifestent les post-actions de toutes les actions $u_0, u_1, \ldots, u_{s-1}$ appliquées au système dans le passé. C'est pourquoi l'action de commande u_s doit être définie compte tenu de l'état du système qui porte l'empreinte de toute la « préhistoire » de ses entrées fournies par la chaîne d'action. Ainsi, le vecteur $\vec{u}_{s-1}$ doit intervenir dans u_s.

Cherchons la stratégie Γ_s optimale dans le sens de la minimisation du risque

$$R = M\{W\} = M\left\{\sum_{s=0}^{n} W_s\right\} = \sum_{s=0}^{n} R_s. \tag{6.358}$$

n est ici fixé comme auparavant.

Trouvons d'abord le risque élémentaire conditionnel. Considérant que x_s^* est imposé et que c'est un paramètre, on peut ne pas l'écrire sous une forme explicite. On a

$$r_s = M\{W_s \mid \vec{w}_{s-1}\} = \int\limits_{\Omega(\bar{\mu}, \vec{u}_s)} W_s[s, F_0, (\vec{z}_s, \vec{u}_s)] \cdot P(\bar{\mu}, \vec{u}_s \mid \vec{w}_{s-1})\, d\Omega. \tag{6.359}$$

La densité de probabilité conjointe de $\bar{\mu}$ et $\vec{u}_s$ est donnée par l'expression

$$P(\bar{\mu}, \vec{u}_s \mid \vec{w}_{s-1}) = P(\bar{\mu} \mid \vec{w}_{s-1}) \cdot P(\vec{u}_s \mid \bar{\mu}, \vec{w}_{s-1}) = \\ = P(\bar{\mu} \mid \vec{w}_{s-1}) \cdot \prod_{i=0}^{s} \Gamma_i (u_i \mid \vec{u}_{i-1}, \vec{w}_{i-1}). \quad (6.360)$$

Cette dernière transformation se vérifie du fait que $\vec{u}_s$ ne dépend que de $\vec{w}_{s-1}$ (cf. fig. 6.21) et l'information sur $\bar{\mu}$ ne change pas la densité de probabilité $P(\vec{u}_s \mid \vec{w}_{s-1})$. De plus, $P(\vec{u}_s \mid \vec{w}_{s-1})$ est évidemment le produit des fonctions Γ_i.

Ensuite, la densité de probabilité conjointe

$$P(\vec{w}_{s-1}, \bar{\mu}) = P(\vec{w}_{s-1} \mid \bar{\mu}) \cdot P(\bar{\mu}) = P(\bar{\mu} \mid \vec{w}_{s-1}) \cdot P(\vec{w}_{s-1}), \quad (6.361)$$

$P(\vec{w}_{s-1})$ étant la densité de probabilité inconditionnelle $\vec{w}_{s-1}$. C'est pourquoi la densité de probabilité a posteriori

$$P_s(\bar{\mu}) = P(\bar{\mu} \mid \vec{w}_{s-1}) = \frac{P(\bar{\mu}) \cdot P(\vec{w}_{s-1} \mid \bar{\mu})}{P(\vec{w}_{s-1})}. \quad (6.362)$$

En portant la formule de Bayes (6.362) dans (6.360) et en tenant compte du fait que

$$P(\vec{w}_{s-1} \mid \bar{\mu}) = \prod_{i=0}^{s-1} P(w_i \mid i, \bar{\mu}), \quad (6.363)$$

on aboutit à l'expression

$$P(\bar{\mu}, \vec{u}_s \mid \vec{w}_{s-1}) = \frac{P(\bar{\mu}) \cdot \prod_{i=0}^{s-1} P(w_i \mid i, \bar{\mu})}{P(\vec{w}_{s-1})} \cdot \prod_{i=0}^{s} \Gamma_i. \quad (6.364)$$

Cette expression peut être portée dans (6.359).

Lorsque le nombre d'expériences est très grand, d'après (6.359) et (6.364) le risque élémentaire s'exprime par la formule

$$R_s = \int_{\Omega(\vec{w}_{s-1})} r_s \cdot P(\vec{w}_{s-1})\, d\Omega = \\ = \int_{\Omega(\bar{\mu}, \vec{u}_s, \vec{w}_{s-1})} W_s(s, \bar{\mu}, \vec{u}_s) \cdot P(\bar{\mu}) \prod_{i=0}^{s-1} P(w_i \mid i, \bar{\mu}) \times \\ \times \prod_{i=0}^{s} \Gamma_i (u_i \mid \vec{u}_{i-1}, \vec{w}_{i-1})\, d\Omega. \quad (6.365)$$

Maintenant la stratégie optimale peut être déterminée de la même façon que dans la théorie de la commande duale. Introduisons la fonction auxiliaire

$$\alpha_k(\vec{u}_k, \vec{w}_{k-1}) = \int\limits_{\Omega(\bar{\mu})} W_k(k, \bar{\mu}, \vec{u}_k) \cdot P(\bar{\mu}) \times \prod_{i=0}^{k-1} P(w_i \mid i, \bar{\mu})\, d\Omega. \qquad (6.366)$$

Les mêmes raisonnements que ceux de la théorie de la commande duale conduisent dans ce cas à un résultat analogue. Il se trouve que la stratégie optimale Γ^*_{n-k} est une stratégie pure. Pour l'obtenir, il faut composer la fonction

$$\gamma_{n-k} = \alpha_{n-k} + \int\limits_{\Omega(w_{n-k})} \gamma^*_{n-k+1}\, d\Omega, \qquad (6.367)$$

de plus $\gamma_n = \alpha_n$ et

$$\gamma^*_i = \min_{u_i \in \Omega(u_i)} \gamma_i = \gamma_i(u^*_i). \qquad (6.368)$$

En minimisant γ_{n-k} par rapport à u_{n-k}, on obtient la commande optimale

$$u^*_{n-k} = u^*_{n-k}(\vec{u}_{n-k-1}, \vec{w}_{n-k-1}). \qquad (6.369)$$

La commande optimale u^*_{n-k} dépend de $\vec{w}_{n-k-1}$ car l'observation de toutes les valeurs précédentes de w_i ($i = 0, 1, \ldots, n-k-1$) permet d'évaluer le vecteur des paramètres $\bar{\mu}$. L'ignorance de l'une quelconque des valeurs de w_i ($i < n-k$) observées antérieurement diminuerait l'information disponible sur $\bar{\mu}$ et, par conséquent, renforcerait l'indétermination de l'estimation des paramètres du système, ce qui peut altérer la commande. Par ailleurs, u^*_{n-k} dépend de $\vec{u}_{n-k-1}$ du fait que toutes les commandes antérieures u_i ($i = 0, 1, \ldots, n-k-1$) laissent leurs « traces » sous la forme de variation de l'état de l'objet B, et donc il faut en tenir compte pour définir la commande optimale à l'instant $t = n-k$.

Ainsi, il se trouve que formellement les opérations de définition de la stratégie optimale du schéma en boucle fermée 6.19 et du schéma en boucle ouverte 6.21 sont quasi identiques.

Mais si l'on regarde ces formules de plus près, on voit que la différence entre elles est bien nette. En effet, comparons les formules de α_k (6.353) et (6.366) pour les systèmes en boucle fermée et en boucle ouverte respectivement. Nous voyons que dans (6.353), la dépendance de $\vec{u}_k$ est en principe plus complexe que dans (6.366). Les coordonnées de ce vecteur font partie non seulement du facteur $W_k(k, \bar{\mu}, \vec{u}_k)$ de l'expression sous le signe somme de (6.353), mais aussi des

facteurs $P(y_i \mid i, \bar{\mu}, \vec{u}_i)$. Cette circonstance reflète le caractère varié de l'influence exercée par la commande u_k sur le risque élémentaire R_s avec $s > k$. D'une part, à l'instant $t = s$, l'action u_k laisse une « trace » dans le système dynamique de l'objet. Pour cette raison W_s dépend du vecteur $\vec{u}_s$ tout entier et non du seul scalaire u_s. De plus, la probabilité a posteriori $P_s(\bar{\mu})$ varie également suivant que la commande u_k ait été bonne ou mauvaise dans le sens d'amélioration des données du système. C'est pourquoi les facteurs $P(y_i, \mid i, \bar{\mu}, \vec{u}_i)$ font partie de l'expression (6.353). Ce deuxième canal d'influence de u_k sur R_s est inexistant dans le système en boucle ouverte. Dans la formule (6.366) de α_k d'un système en boucle ouverte, les facteurs $P(w_i \mid i, \bar{\mu})$ sont indépendants de $\vec{u}_i$.

Toutefois, il existe des cas où le système 6.19 devient réductible. Cela est possible lorsque la formule (6.353) peut être mise sous une forme analogue à celle de (6.366). Si l'on peut trouver une fonction

$$w_i = w_i(y_i, i, \vec{u}_i) \tag{6.370}$$

telle qu'elle vérifie l'égalité

$$P(y_i \mid i, \bar{\mu}, \vec{u}_i) = P(w_i \mid i, \bar{\mu}), \tag{6.371}$$

(6.353) se ramène alors à (6.366) et le système en boucle fermée devient réductible. Il faut en outre que soit vérifiée la condition

$$\int\limits_{\Omega(w_i)} P(w_i \mid i, \bar{\mu})\, d\Omega = 1, \tag{6.372}$$

car l'expression sous le signe somme de (6.372) est une densité de probabilité de w_i.

Un exemple de système réductible à objet avec inertie a été examiné dans ce qui précède [cf. équation (6.222) et la suite]. La formule (6.259) montre que dans ce système la mesure de $\vec{x}_i$ et $\vec{u}_i$ est équivalente à la mesure de μ avec une erreur. Pourtant, la classe des systèmes réductibles est très étroite par rapport à celles des systèmes irréductibles, pour lesquels les conditions de types (6.370) et (6.371) ne sont pas valables.

L'existence des systèmes irréductibles est une illustration de plus du fait que, dans un certain sens, la notion de chaîne de réaction a une valeur absolue. Les processus dont les systèmes asservis sont le siège sont plus riches de contenu que les processus se déroulant dans des systèmes en boucle ouverte ; les systèmes à réaction se caractérisent, entre autres, par des traits inexistants dans les systèmes en boucle ouverte.

Nous étudions dans ce qui suit certains problèmes relatifs à la construction des schémas fonctionnels des organes de commande faisant partie des systèmes irréductibles. Considérons, à titre d'illustration, l'exemple très simple d'un système irréductible au schéma fonctionnel 6.22. Ce schéma ne diffère du schéma apparenté 6.7,

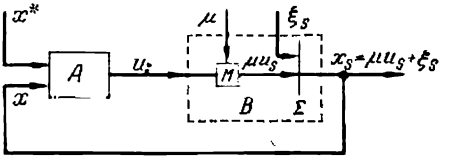


Fig. 6.22.

où $x_s = \mu u_s$, que par le point d'application d'un bruit purement aléatoire. Au lieu des variables indépendantes h_s dans la chaîne de réaction (fig. 6.6), le schéma 6.22 comporte des variables indépendantes ξ_s à l'intérieur du système B. Supposons que ces variables soient gaussiennes de valeur moyenne et de densité de probabilité identique

$$P(\xi_s) = q(\xi_s) = \frac{1}{\sigma_\xi \sqrt{2\pi}} \exp\left\{-\frac{\xi_s^2}{2\sigma_\xi^2}\right\}. \tag{6.373}$$

Le schéma fonctionnel de l'objet B comprend un multiplieur M et un additionneur Σ. Par conséquent,

$$x_s = \mu u_s + \xi_s. \tag{6.374}$$

Le « gain » μ du multiplieur M dont la grandeur de sortie est égale à μu_s est une variable munie d'une loi normale et à densité de probabilité a priori

$$P_0(\mu) = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right\}. \tag{6.375}$$

La fonction de pertes élémentaire s'écrit

$$W_s = (x_s^* - x_s)^2, \tag{6.376}$$

où x_s^* est la grandeur constante imposée, généralement différente du zéro.

Si le bruit aléatoire ξ_s n'existait pas ($\xi_s = 0$), une seule mesure des grandeurs u_s et x_s suffirait pour connaître la valeur de μ. Donc dans ce cas la définition de la caractéristique du système B constituerait un problème trivial. D'autre part, si la grandeur μ est connue et si $\xi_s \neq 0$, la caractéristique du système B subit d'un pas à l'autre des variations aléatoires. Toutefois, l'accumulation de l'information sur la caractéristique du système est en principe impossible du fait que, pour $i \neq j$, certaines valeurs de ξ_i et ξ_j sont indépendantes.

Aussi, l'étude des caractéristiques du système n'a-t-elle aucun sens. Le problème ne devient réel que si la grandeur μ est inconnue et, de plus, si $\xi_s \neq 0$. L'obtention d'une mesure précise de μ en un nombre fini de pas est alors impossible. En accumulant l'information sur la mesure des u_i et x_i on peut construire à chaque nouveau pas de nouvelles densités de probabilité a posteriori $P_{i+1}(\mu)$ qui permettent d'obtenir des estimations de plus en plus précises de μ en contribuant ainsi à une organisation toujours meilleure de la commande.

Dans les chapitres précédents nous n'avons pas examiné les problèmes sur les systèmes à bruit de fond ξ_s. Or, ce sont des problèmes qui autorisent l'application sans modification, tant soit peu tangible, de la théorie de la commande duale. Déduisons d'abord la formule du risque élémentaire conditionnel r_s. La grandeur x_s^* étant donnée, il n'est pas de rigueur de l'écrire sous une forme explicite. Le schéma de la figure 6.22 et la formule (6.376) donnent

$$r_s = M\{W_s \mid \vec{x}_{s-1}, \vec{u}_{s-1}\} = \\ = \int\limits_{\Omega(x_s,\, u_s)} (x_s^* - x_s)^2 P_s(x_s \mid u_s) \cdot \Gamma_s(u_s \mid \vec{x}_{s-1}, \vec{u}_{s-1})\, d\Omega. \qquad (6.377)$$

Ici $P_s(x_s \mid u_s)$ est la densité de probabilité a posteriori de x_s, u_s étant fixée. Cette fonction caractérise l'information sur le système, disponible avant le s-ième pas. De même que dans ce qui précède, l'expression $\Gamma_s(u_s \mid \vec{x}_{s-1}, \vec{u}_{s-1})$ est utilisée pour désigner une stratégie aléatoire de l'organe de commande A.

La fonction $P_s(x_s \mid u_s)$ est associée à la densité de probabilité a posteriori $P_s(\mu)$ de la variable μ par la relation évidente

$$P_s(x_s \mid u_s) = \int\limits_{\Omega(\mu)} P(x_s \mid \mu, u_s)\, P_s(\mu)\, d\Omega, \qquad (6.378)$$

où $P(x_s \mid \mu, u_s)$ est la densité de probabilité conditionnelle de x_s, μ et u_s étant fixées, et $\Omega(\mu)$ est le domaine des valeurs possibles de μ.

En vertu de (6.374) on a

$$\xi_s = x_s - \mu u_s. \qquad (6.379)$$

Par conséquent, d'après (6.373)

$$P(x_s \mid \mu, u_s) = \frac{1}{\sigma_\xi \sqrt{2\pi}} \exp\left\{-\frac{(x_s - \mu u_s)^2}{2\sigma_\xi^2}\right\}. \qquad (6.380)$$

Tel est ainsi le premier facteur de l'expression sous le signe somme de (6.378).

Cherchons maintenant à l'aide de la formule de Bayes l'expression $P_s(\mu)$ qui est le deuxième facteur de (6.378) :

$$P_s(\mu) = P(\mu \mid \vec{u}_{s-1}, \vec{x}_{s-1}) = \frac{P(\vec{u}_{s-1}, \vec{x}_{s-1} \mid \mu)\, P_0(\mu)}{P(\vec{u}_{s-1}, \vec{x}_{s-1})} \qquad (6.381)$$

Considérons l'expression

$$P(u_i, x_i \mid \mu, \vec{u}_{i-1}, \vec{x}_{i-1}) = \\ = P(x_i \mid \mu, u_i, \vec{u}_{i-1}, \vec{x}_{i-1})\, P(u_i \mid \mu, \vec{u}_{i-1}, \vec{x}_{i-1}) = \\ = P(x_i \mid \mu, u_i) \cdot \Gamma_i (u_i \mid \vec{u}_{i-1}, \vec{x}_{i-1}). \qquad (6.382)$$

Nous avons rejeté $\vec{u}_{i-1}$, $\vec{x}_{i-1}$ du premier facteur du second membre puisque la fixation de μ et de u_i définit complètement la densité de probabilité conditionnelle de x_i et l'addition de $\vec{u}_{i-1}$ et $\vec{x}_{i-1}$ ne fournit aucune information nouvelle du fait que le système B est sans mémoire. Le deuxième facteur est la stratégie Γ_i de l'organe de commande A qui n'est assujetti qu'aux $\vec{u}_{i-1}$ et $\vec{x}_{i-1}$. C'est pourquoi on peut rejeter μ à droite du trait vertical de ce facteur.

(6.382) entraîne

$$P(\vec{u}_{s-1}, \vec{x}_{s-1} \mid \mu) = \prod_{i=0}^{s-1} P(x_i \mid \mu, u_i) \cdot \prod_{i=0}^{s-1} \Gamma_i. \qquad (6.383)$$

Par ailleurs, le risque élémentaire moyen

$$R_s = \int\limits_{\Omega(\vec{u}_{s-1}, \vec{x}_{s-1})} r_s P(\vec{u}_{s-1}, \vec{x}_{s-1})\, d\Omega.$$

En portant (6.383) dans (6.381), ensuite (6.381) et (6.380) dans (6.378), puis (6.378) dans (6.377) et (6.377) dans l'expression de R_s, on obtient, compte tenu de (6.375), à un coefficient de proportionnalité près,

$$R_s \equiv \int\limits_{\Omega(\mu, \vec{u}_s, \vec{x}_s)} (x_s^* - x_s)^2 \exp\left\{-\frac{1}{2\sigma_\xi^2} \sum_{i=0}^{s} (x_i - \mu u_i)^2 - \right. \\ \left. - \frac{(\mu - \mu_0)^2}{2\sigma_\mu^2}\right\} \cdot \prod_{i=0}^{s} \Gamma_i\, d\Omega. \qquad (6.384)$$

Dans le cas considéré, la méthode générale de recherche est la même que dans ceux examinés plus haut.

Introduisons la fonction

$$\alpha_s(u_s, \vec{x}_{s-1}, \vec{u}_{s-1}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_s^* - x_s)^2 \times \\ \times \exp\left\{-\frac{1}{2\sigma_\xi^2} \sum_{i=0} (x_i - \mu u_i)^2 - \frac{(\mu - \mu_0)^2}{2\sigma_\mu^2}\right\} dx_s\, d\mu. \qquad (6.385)$$

Définissons ensuite u_n^* en minimisant α_n par rapport à u_n, puis trouvons γ_{n-1}. En minimisant γ_{n-1} par rapport à u_{n-1} on trouve u_{n-1}^*, etc.

La fonction α_s peut être établie sous sa forme finie. Après l'intégration de (6.385) à un coefficient de proportionnalité près on a

$$\alpha_s \equiv \exp\left\{\frac{1}{2\rho_\xi^2}\sum_{i=0}^{s-1} x_i^2\right\}\left[\sigma_\xi^2+(x_s^*)^2+\frac{u_s^2}{2r_{s-1}}\left(1+\frac{2q_{s-1}^2}{r_{s-1}}\right)-\right.$$

$$\left.-2u_s x_s^*\frac{q_{s-1}}{r_{s-1}}\right]\times\frac{1}{\sqrt{r_{s-1}}}\exp\left\{\frac{q_{s-1}^2}{r_{s-1}}\right\}, \qquad (6.386)$$

où

$$\left.\begin{aligned} q_{s-1} &= \frac{1}{2\sigma_\xi^2}\sum_{i=0}^{s-1} x_i u_i+\frac{\mu_0}{2\sigma_\mu^2}, \\ r_{s-1} &= \frac{1}{2\sigma_\xi^2}\sum_{i=0}^{s-1} u_i^2+\frac{1}{2\sigma_\mu^2}. \end{aligned}\right\} \qquad (6.387)$$

La fonction α_n est minimisée lorsque pour $s = n$ dans l'expression entre crochets de (6.386) la dérivée par rapport à u_n s'annule:

$$\frac{u_n}{r_{n-1}}\left(1+\frac{2q_{n-1}^2}{r_{n-1}}\right)-2x_n^*\frac{q_{n-1}}{r_{n-1}}=0. \qquad (6.388)$$

On en déduit que

$$u_n^*=2x_n^*\frac{q_{n-1}}{1+\dfrac{2q_{n-1}^2}{r_{n-1}}}. \qquad (6.389)$$

Ainsi, la meilleure valeur de u_n^* est définie par l'expression $\mu' u_n^* = x_n^*$, où μ' est la «meilleure» estimation de la grandeur μ. (6.389) implique

$$\mu'=\frac{x_n^*}{u_n^*}=\frac{1}{2q_{n-1}}\left(1+\frac{2q_{n-1}^2}{r_{n-1}}\right)=\frac{1}{2q_{n-1}}+\frac{q_{n-1}}{r_{n-1}}=$$

$$=\frac{1}{\dfrac{\mu_0}{\sigma_\mu^2}+\dfrac{1}{\sigma_\xi^2}\displaystyle\sum_{i=0}^{n-1} x_i u_i}+\frac{\dfrac{1}{2\sigma_\xi^2}\displaystyle\sum_{i=0}^{n-1} x_i u_i+\dfrac{\mu_0}{2\sigma_\mu^2}}{\dfrac{1}{2\sigma_\xi^2}\displaystyle\sum_{i=0}^{n-1} u_i^2+\dfrac{1}{2\sigma_\mu^2}}. \qquad (6.390)$$

Dans le cas d'un grand nombre d'essais, lorsque les sommes $\sum x_i u_i$ et $\sum u_i^2$ deviennent grandes, cette estimation s'approche de la

grandeur

$$\mu'' = \frac{\sum_{i=0}^{n-1} x_i u_i}{\sum_{i=0}^{n-1} u_i^2}. \tag{6.391}$$

Si la relation $x_i = \mu u_i$ était vérifiée, (6.391) conduirait à l'égalité $\mu'' = \mu$.

La grandeur γ_{n-1} est définie par l'expression

$$\gamma_{n-1} = \alpha_{n-1} + \int_{-\infty}^{\infty} \alpha_n^* \, dx_{n-1}. \tag{6.392}$$

Le résultat de l'intégration ne se traduit pas par une formule finie. Pourtant, la chaîne des intégrations et des minimisations peut se faire en ne retenant que les fonctions de deux variables, de même que dans l'exemple apparenté que nous avons examiné dans ce qui précède [cf. les équations (6.141) et (6.145)]. Comme le. montre (6.387), les variables correspondantes sont les grandeurs

$$L_{s-1} = \sum_{i=0}^{s-1} x_i u_i, \quad M_{s-1} = \sum_{i=0}^{s-1} u_i^2. \tag{6.393}$$

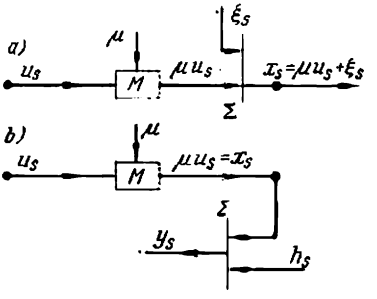


Fig. 6.23.

De même que dans l'exemple donné plus haut, dans le cas considéré la commande optimale u_s^* est associée aux valeurs L_{s-1} et M_{s-1}. Il est utile également d'insister sur le fait que u_s^* dépend en général du numéro du pas ou, plus précisément, du nombre de pas $(n - s)$ à faire jusqu'à la fin du processus. Ainsi

$$u_s^* = u_s^* [L_{s-1}, M_{s-1}, (n-s)]. \tag{6.394}$$

Il existe également une autre méthode pour justifier le choix de L_{s-1} et M_{s-1} en tant que caractéristiques suffisantes ou d'autres grandeurs qui leur sont associées. Examinons le système commandé B de la figure 6.23,a. Soit ξ_s toujours le processus strictement aléatoire caractérisé par la densité de probabilité $q(\xi_s)$ définie par l'équation (6.373). Supposons que la densité de probabilité a posteriori $P_{s-1}(\mu)$ de la grandeur μ est définie par l'expression

$$P_{s-1}(\mu) = \frac{1}{\sigma_{s-1}\sqrt{2\pi}} \exp\left\{-\frac{(\mu - m_{s-1})^2}{2\sigma_{s-1}^2}\right\}, \tag{6.395}$$

m_{s-1} et σ_{s-1} étant respectivement la valeur moyenne et l'écart-type. Supposons encore que le système subit à l'entrée une certaine action u_s qui entraîne à la sortie une certaine réponse x_s. Une question se pose : quelle est la nouvelle densité a posteriori $P_s(\mu)$ trouvée compte tenu de la nouvelle observation effectuée au s-ième pas? Il est évident que $P_s(\mu)$ doit dépendre des valeurs σ_{s-1}, m_{s-1} caractérisant l'ancienne densité $P_{s-1}(\mu)$ qui dans ce cas joue le rôle d'une densité a priori, ainsi que des valeurs de u_s et x_s :

$$P_s(\mu) = P_s(\mu, \sigma_{s-1}, m_{s-1}, u_s, x_s) = P_s(\mu \mid u_s, x_s, \sigma_{s-1}, m_{s-1}). \quad (6.396)$$

La forme explicite de l'expression de $P_s(\mu)$ se trouve facilement. En représentant $P(x_s, \mu \mid u_s)$ sous la forme d'un produit, on obtient sans difficulté la formule de Bayes

$$P_s(\mu) = P(\mu \mid x_s, u_s) = \frac{P_{s-1}(\mu)\, P(x_s \mid \mu, u_s)}{P(x_s)} \quad (6.397)$$

$P(x_s)$ étant une certaine grandeur indépendante de μ et qui de ce fait n'intervient pas en principe dans la formule (6.397), on peut la poser égale à la constante C_1. La grandeur de cette constante se calcule d'après la condition

$$\int_{-\infty}^{\infty} P_s(\mu)\, d\mu = 1. \quad (6.398)$$

En remplaçant $P_{s-1}(\mu)$ par son expression tirée de (6.395) et $P(x_s \mid \mu, u_s)$ par son expression tirée de (6.380), on a

$$P_s(\mu) = C_2 \exp\left\{-\frac{(x_s - \mu u_s)^2}{2\sigma_\xi^2} - \frac{(\mu - m_{s-1})^2}{2\sigma_{s-1}^2}\right\}, \quad (6.399)$$

où C_2 est une constante. Cette formule se met facilement sous la forme

$$P_s(\mu) = C_3 \exp\left\{-\frac{(\mu - m_s)^2}{2\sigma_s^2}\right\}. \quad (6.400)$$

où C_3 est une nouvelle constante, alors que m_s et σ_s sont définies par les expressions

$$\left.\begin{aligned} m_s &= \frac{u_s x_s + \left(\frac{\sigma_\xi}{\sigma_{s-1}}\right)^2 m_{s-1}}{u_s^2 + \left(\frac{\sigma_\xi}{\sigma_{s-1}}\right)^2}, \\ \sigma_s^2 &= \frac{\sigma_{s-1}^2}{1 + \left(\frac{\sigma_{s-1}}{\sigma_\xi}\right)^2 u_s^2}. \end{aligned}\right\} \quad (6.401)$$

D'après les conditions de normalisation il est clair que $C_3 = \dfrac{1}{\sigma_s \sqrt{2\pi}}$.

Les formules (6.400) et (6.401) permettent de tirer plusieurs conclusions utiles. Premièrement, il se trouve que la répartition a posteriori caractérisée par la formule de $P_s(\mu)$ est également normale. On en déduit que lorsque la densité de probabilité a priori $P_0(\mu)$ est imposée sous la forme d'une loi normale (6.375), les densités de probabilité suivantes $P_1(\mu)$, puis $P_2(\mu)$, $P_3(\mu)$, etc., sont également décrites par des lois normales. Par conséquent, deux nombres seulement suffisent pour caractériser une densité de probabilité a posteriori quelconque, celui de la valeur moyenne m_s et de la variance σ_2^s. Ensuite, l'expression (6.401) montre que la nouvelle moyenne m_s s'établit à partir de l'ancienne valeur m_{s-1} par transformation linéaire dont font partie, en plus de σ_{s-1}, les fonctions des valeurs de u_s et x_s obtenues au s-ième pas : u_s^2 et $u_s x_s$. Enfin, la nouvelle variance σ_s^2 est inférieure à la variance ancienne σ_{s-1}^2 et, de plus, σ_s^2 dépend de la valeur u_s^2 (mais non pas de la valeur x_s).

Le schéma de la figure 6.23,*a* n'est pas le seul à jouir de la propriété de conserver la loi normale. Considérons, par exemple, un autre schéma du système commandé B (cf. fig. 6.23,*b*). Ici le « bruit de fond » ξ_s n'existe pas ; pourtant, la mesure de la grandeur de sortie x_s s'effectue avec une erreur h_s. Pour cette raison on introduit dans l'organe de commande A la grandeur

$$y_s = x_s + h_s = \mu u_s + h_s. \tag{6.402}$$

Soit $P_{s-1}(\mu)$ la densité de probabilité a posteriori de μ après le $(s-1)$-ième pas. Cherchons la densité de probabilité a posteriori suivante $P_s(\mu)$ une fois que les valeurs de u_s et y_s seront fixées au s-ième pas. Il est clair que l'égalité analogue à (6.397) est vraie :

$$P_s(\mu) = P(\mu \mid y_s, u_s) = \frac{P_{s-1}(\mu)\, P(y_s \mid \mu, u_s)}{P(y_s)}. \tag{6.403}$$

Supposons que $P_{s-1}(\mu)$ soit décrite par la loi normale (6.395). D'après la formule (6.402) et la loi de répartition (6.120) de h_s on a

$$P(y_s \mid \mu, u_s) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{(y_s - \mu u_s)^2}{2\sigma_h^2}\right\}. \tag{6.404}$$

En portant (6.395) et (6.404) dans (6.403), on amène

$$P_s(\mu) = C_2 \exp\left\{-\frac{(y_s - \mu u_s)^2}{2\sigma_h^2} - \frac{(\mu - m_{s-1})^2}{2\sigma_{s-1}^2}\right\} =$$

$$= C_3 \exp\left\{-\frac{(\mu - m_s)^2}{2\sigma_s^2}\right\}, \tag{6.405}$$

où C_2 et C_3 sont des constantes. L'expression (6.405) ne se distingue de (6.399) que par le fait que x_s y est remplacée par y_s et σ_ξ par σ_h. C'est pourquoi les formules obtenues de m_s et σ_s sont

analogues à (6.401) :

$$\left.\begin{aligned} m_s &= \frac{u_s y_s + \left(\frac{\sigma_h}{\sigma_{s-1}}\right)^2 m_{s-1}}{u_s^2 + \left(\frac{\sigma_h}{\sigma_{s-1}}\right)^2}, \\ \sigma_s^2 &= \frac{\sigma_{s-1}^2}{1 + \left(\frac{\sigma_{s-1}}{\sigma_h}\right)^2 u_s^2}. \end{aligned}\right\} \tag{6.406}$$

A tout instant $t = s$ la densité de probabilité disponible $P_{s-1}(\mu)$ peut être considérée comme une densité a priori. Par conséquent, pour $t = s$ la commande optimale u_s^* doit être choisie en fonction des paramètres m_{s-1} et σ_{s-1} de la fonction $P_{s-1}(\mu)$, ainsi qu'en fonction du nombre de pas à faire jusqu'à la fin du processus. Ce dernier nombre est égal à $(n - s)$. De plus, u_s^* dépend de x_s^*. Ainsi,

$$u_s^* = u_s^* [x_s^*, \ m_{s-1}, \ \sigma_{s-1}, \ (n-s)]. \tag{6.407}$$

En partant des formules (6.401) et (6.407) on peut construire le schéma fonctionnel d'un organe de commande A pour le système représenté sur la figure 6.23,*a*.

Lors du passage de $P_{s-1}(\mu)$ à $P_s(\mu)$ le caractère de la loi de répartition peut se conserver même si la répartition est régie par d'autres lois que la loi normale (cf. [5.32, 5.33, 6.16]). Soit, dans un cas plus général, la densité de probabilité P_{s-1} du vecteur $\bar{\mu}$ qui est une fonction des paramètres $a_{1,\,s-1}, \ldots, a_{q,\,s-1}$ réunis en un vecteur $\bar{a}_{s-1}$:

$$P_{s-1}(\mu) = P_{s-1}(\bar{\mu}, a_{1,\,s-1}, \ldots, a_{q,\,s-1}) = P_{s-1}(\bar{\mu}, \bar{a}_{s-1}). \tag{6.408}$$

Supposons ensuite que l'observation des valeurs de $\bar{u}_s$ et $\bar{y}_s$ à la sortie de l'organe de commande A et à son entrée liée à la sortie du système commandé permet d'obtenir au lieu du vecteur $\bar{a}_{s-1}$ un nouveau vecteur $\bar{a}_s$. Ce vecteur est fourni par une transformation $\bar{\Phi}$ de l'ancien vecteur $\bar{a}_{s-1}$ et dépend certainement aussi de $\bar{u}_s$ et $\bar{y}_s$.

$$\bar{a}_s = \bar{\Phi}(\bar{a}_{s-1}, \bar{u}_s, \bar{y}_s). \tag{6.409}$$

Les expressions (6.406) sont un cas particulier de cette relation.

Le vecteur $\bar{a}_s$ peut s'appeler vecteur d'information puisqu'il résume toute l'information a posteriori sur les caractéristiques du système commandé (mais non pas sur son état). La fonction $P_s(\bar{\mu})$ dépend de ce vecteur

$$P_s(\bar{\mu}) = P_s(\bar{\mu}, \bar{a}_s). \tag{6.410}$$

En considérant que $P_{s-1}(\bar{\mu})$ est une densité de probabilité a priori avant le s-ième pas, on voit sans peine que la commande optimale $\bar{u}_s^*$

est la fonction du vecteur $\bar{a}_{s-1}$, et dans le cas d'un système avec inertie, elle est également la fonction des coordonnées de ce vecteur, c'est-à-dire du vecteur $\bar{x}_{s-1}$. (Si la valeur exacte de $\bar{x}_{s-1}$ ne peut être mesurée, son estimation est donnée par les vecteurs $\vec{y}_{s-1}$ et $\vec{u}_{s-1}$.) En outre, u_s^* dépend de $\bar{x}_s^*$ et du nombre de pas $(n-s)$ à faire jusqu'à la fin du processus. Ainsi, dans le cas particulier de $\bar{h}_s = 0$ et $\bar{x}_s = \bar{y}_s$ l'expression vraie est

$$\bar{u}_s^* = \bar{u}_s^* [\bar{a}_{s-1}, \bar{x}_{s-1}, \bar{x}_s^*, (n-s)]. \tag{6.411}$$

Les composantes des vecteurs $\bar{a}_{s-1}$ et $\bar{x}_{s-1}$ qui traduisent l'information sur les caractéristiques et l'état du système peuvent être appelées dans ce cas coordonnées suffisantes. L'application des coordonnées suffisantes était recommandée pour résoudre les problèmes

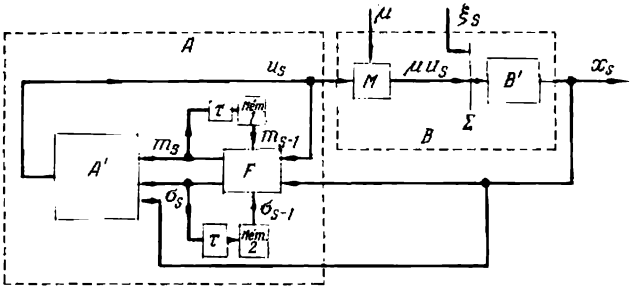


Fig. 6.24.

statistiques de commande optimale qui ne se rapportent pas au domaine de la théorie de la commande duale exposée dans [5.33, 5.40 et 6.10]. Les méthodes exposées dans [6.16] relatives aux coordonnées suffisantes sont appliquées pour résoudre un exemple décrit ci-dessous de la théorie de commande duale.

La grandeur aléatoire μ et la suite des grandeurs indépendantes ξ_s sont régies de même que dans ce qui précède par les lois de répartition (6.375) et (6.373) respectivement (fig. 6.24). Le sous-bloc B' est un intégrateur discret à grandeur d'entrée $\mu u_s + \xi_s$, l'équation du système B tout entier s'écrivant

$$x_{s+1} - x_s = \mu u_s + \xi_s. \tag{6.412}$$

Si la densité de probabilité a priori $P_s(\mu)$ de μ est régie par une loi normale de moyenne m_s et de variance σ_s^2, la densité a posteriori obtenue après la mesure des valeurs de u_s, x_s et x_{s+1} est régie également par une loi normale de moyenne m_{s+1} et de variance σ_{s+1}^2. On le montre facilement de la même manière que dans les exemples

exposés plus haut. En effet, si l'on considère $P_s(\mu)$ comme une densité de probabilité « a priori » de μ, on vérifie alors l'égalité

$$P(\mu, x_{s+1} \mid x_s, u_s) = P_s(\mu) \cdot P(x_{s+1} \mid x_s, \mu, u_s) = \\ = P(x_{s+1} \mid u_s, x_s) \cdot P(\mu \mid x_s, x_{s+1}, u_s). \quad (6.413)$$

$P(x_{s+1} \mid x_s, u_s)$ étant indépendante de μ, remplaçons cette densité de probabilité par $(C_1)^{-1}$. Il vient

$$P_{s+1}(\mu) = P(\mu \mid x_s, x_{s+1}, u_s) = \frac{P_s(\mu) P(x_{s+1} \mid x_s, u_s, \mu)}{P(x_{s+1} \mid x_s, u_s)} = \\ = C_1 \cdot P_s(\mu) \cdot P(x_{s+1} \mid x_s, \mu, u_s). \quad (6.414)$$

D'après (6.412) et (6.373)

$$P(x_{s+1} \mid \mu, x_s, u_s) = \frac{1}{\sigma_\xi \sqrt{2\pi}} \exp\left\{-\frac{(x_{s+1} - x_s - \mu u_s)^2}{2\sigma_\xi^2}\right\}. \quad (6.415)$$

En portant cette expression dans (6.414) avec la formule de $P_s(\mu)$:

$$P_s(\mu) = \frac{1}{\sigma_s \sqrt{2\pi}} \exp\left\{-\frac{(\mu - m_s)^2}{2\sigma_s^2}\right\}, \quad (6.416)$$

on aboutit, après des transformations peu compliquées à l'expression

$$P_{s+1}(\mu) = C_2 \exp\left\{-\frac{(\mu - m_{s+1})^2}{2\sigma_{s+1}^2}\right\}, \quad (6.417)$$

où C_2 est une constante, alors que m_{s+1} et σ_{s+1} sont fournis par les formules

$$\left.\begin{aligned} m_{s+1} &= \frac{u_s(x_{s+1} - x_s) + \left(\frac{\sigma_\xi}{\sigma_s}\right)^2 m_s}{u_s^2 + \left(\frac{\sigma_\xi}{\sigma_s}\right)^2}, \\ \sigma_{s+1}^2 &= \frac{\sigma_s^2}{1 + \left(\frac{\sigma_s}{\sigma_\xi}\right)^2 u_s^2}. \end{aligned}\right\} \quad (6.418)$$

La figure 6.24 représente le bloc Φ régi par l'équation (6.418).

Comptons maintenant les pas à rebours, à partir de la fin du processus. Ainsi, les u_1, x_1 seront les dernières valeurs, les u_2, x_2 les avant-dernières, etc.

Soit la fonction de pertes de la forme

$$W_s = x_{s-1}^2 + u_s^2. \quad (6.419)$$

W_s est donc définie par l'équation de u_s et celle des sorties x_{s-1} qui apparaîtra au pas suivant.

Introduisons la notation

$$S_r = \sum_{i=1}^{r} R_i = \sum_{i=1}^{r} M\{W_i\} = \sum_{i=1}^{r} M\{x_{i-1}^2 + u_i^2\}. \qquad (6.420)$$

Convenons de désigner par S_r^* la valeur optimale de la fonction S_r obtenue dans les conditions des commandes optimales $u_r^*, u_{r-1}^*, \ldots$ $\ldots, u_1^*$ appartenant au domaine admissible $\Omega(u_i)$:

$$S_r^* = \min S_r = \min \sum_{i=1}^{r} M\{x_{i-1}^2 + u_i^2\}. \qquad (6.421)$$

Les méthodes utilisées dans [6.16] pour établir les formules de la commande optimale relatives à l'exemple considéré se distinguent de celles que nous avons adoptées dans ce chapitre, bien qu'elles leur soient tofent à fait équivalentes et se basent également sur la programmation dynamique. Dans ces conditions, une circonstance importante est que la fonction S_r^* dépend de l'état « initial » du système B (pour un système d'ordre un cet état est défini par une seule coordonnée x_r), des paramètres m_r, σ_r de densité de probabilité «initiale » $P_r(\mu)$ de μ, ainsi que du nombre r de pas à faire jusqu'à la fin du processus. Ainsi

$$S_r^* = S_r^*(r, x_r, m_r, \sigma_r) = S_r^*(x_r, m_r, \sigma_r).$$

La lettre r peut être omise entre les parenthèses, car elle figure dans l'indice (S_r^*).

Suivant le principe de la programmation dynamique, de même qu'au chapitre IV, nous avons

$$S_r^*(x_r, m_r, \sigma_r) = \min_{u_r} M\{W_r + \min_{u_{r-1}, \ldots, u_1} M\{\sum_{i=1}^{r-1} W_i\}\} =$$
$$= \min_{u_r} M\{x_{r-1}^2 + u_r^2 + S_{r-1}^*(x_{r-1}, m_{r-1}, \sigma_{r-1}). \} \qquad (6.422)$$

Par ailleurs, l'étude du dernier pas conduit à la formule

$$S_1^*(x_1, m_1, \sigma_1) = \min_{u_1} M\{x_0^2 + u_1^2\}. \qquad (6.423)$$

Ici on entend sous x_0 l'état après l'application lors du dernier pas de la commande u_1, c'est-à-dire l'état à la fin du processus tout entier.

En vertu de (6.412) et avec la nouvelle numérotation des pas on a

$$x_{r-1} = x_r + \mu u_r + \xi_r. \qquad (6.424)$$

En remplaçant dans (6.418) la différence $(x_{s+1} - x_s)$ ou, d'après la nouvelle numérotation, $(x_{r-1} - x_r)$ par la valeur égale $\mu u_r +$

$+ \xi_r$, on tombe sur les expressions

$$\left.\begin{aligned} m_{r-1} &= \frac{\mu u_r^2 + m_r \left(\frac{\sigma_\xi}{\sigma_r}\right)^2 + u_r \xi_r}{u_r^2 + \left(\frac{\sigma_\xi}{\sigma_r}\right)^2}, \\ \sigma_{r-1}^2 &= \frac{\sigma_r^2}{1 + \left(\frac{\sigma_r}{\sigma_\xi}\right)^2 u_r^2}. \end{aligned}\right\} \tag{6.425}$$

Voyons maintenant s'il est possible d'effectuer une chaîne de calculs des valeurs de u_r^* à partir des équations (6.422)-(6.425). La définition de u_r^* d'après (6.422) impose un calcul préalable de S_{r-1}^*. En outre, les expressions (6.425) doivent figurer comme arguments de S_{r-1}^*. Pourtant, la valeur de μ dans la première d'entre elles est inconnue. C'est pourquoi dans [6.16] on prend la moyenne de cette expression par rapport à μ en considérant que la densité de probabilité a priori de μ est $P_0(\mu)$. Cette opération s'explique par le fait que lorsque le nombre d'épreuves est grand, la formule de répartition de μ a précisément cette forme. Or maintenant S_{r-1}^* est la fonction d'une valeur moyenne de la grandeur aléatoire m_{r-1}. Des formules analogues peuvent se déduire aussi à l'aide des méthodes exposées aux paragraphes précédents de ce chapitre.

La chaîne des calculs commence par la formule (6.423) où x_0 est remplacée par sa valeur tirée de (6.424)

$$\begin{aligned} S_1^*(x_1, m_1, \sigma_1) &= \min_{u_1} M\{x_0^2 + u_1^2\} = \\ &= \min_{u_1} M\{u_1^2 + (x_1 + \mu u_1 + \xi_1)^2\} = \\ &\min_{u_1} [x_1^2 + (m_1^2 + \sigma_1^2) u_1^2 + \sigma_\xi^2 + 2m_1 x_1 u_1 + u_1^2]. \end{aligned} \tag{6.426}$$

Après avoir chassé les parenthèses on cherche la moyenne ; comme nous l'avons indiqué plus haut, nous avons adopté pour μ la répartition normale a priori à moyenne m_1 et à variance σ_1^2. L'obtention de la moyenne de la grandeur ξ_1^2 entraîne l'apparition du terme σ_ξ^2.

En égalant à zéro la dérivée de l'expression entre crochets de (6.426), on obtient u_1^* et puis S_1^* :

$$\left.\begin{aligned} u_1^* &= -\frac{m_1 x_1}{m_1^2 + \sigma_1^2 + 1}. \\ S_1^*(x_1, m_1, \sigma_1) &= \frac{x_1^2 (1 + \sigma_1^2)}{m_1^2 + \sigma_1^2 + 1} + \sigma_\xi^2. \end{aligned}\right\} \tag{6.427}$$

Passons à la définition de S_2^* et u_2^*. (6.422) amène

$$S_2^*(x_2, m_2, \sigma_2) = \min_{u_2} M\{x_1^2 + u_2^2 + S_1^*(x_1, m_1, \sigma_1)\}. \tag{6.428}$$

Si l'on introduit ici l'expression de S_1^* tirée de (6.427) et celle de x_1 tirée de (6.424), on obtient une fonction de u_2 assez complexe et la définition de u_2^* par la méthode analytique devient difficile. C'est pourquoi les calculs ultérieurs ont été effectués sur un calculateur numérique. Lors de l'établissement du programme du problème la fonction S_r^* a été approximée à l'aide des polynômes jusqu'à la puissance quatre. Comme l'indique [6.16], après les calculs et la définition des relations

$$u_i^* = u_i^* (x_i, m_i, \sigma_i) \tag{6.429}$$

il s'est avéré que dès le quatrième pas à partir de la fin, le caractère de ces relations se stabilise et l'approximation par les égalités

$$u_i^* \cong \frac{-m_i x_i}{m_i^2 + \frac{1}{2}\sigma_i^2 + 1} \tag{6.430}$$

leur assure une précision suffisante pour la pratique.

Cette loi peut être rangée dans le sous-bloc A' représenté sur la figure 6.24.

Si la commande s'effectuait sans identification simultanée de l'objet, il est facile de montrer que les valeurs absolues des u_i seraient plus petites. L'étude du système réalisée simultanément avec la commande impose l'augmentation des valeurs de u_i, car plus u_i est grande, plus le produit μu_i est grand par rapport au bruit ξ_i et, par conséquent, plus l'estimation du gain inconnu μ est exacte. Ainsi plus $|u_i|$ est grande, plus le processus d'étude de la caractéristique du système commandé B est intense. Toutefois, la forme de la fonction de pertes W_i (6.419) montre qu'une valeur excédentaire des u_i est à éviter. Premièrement, elle accroît le « coût » de la commande, c'est-à-dire le terme u_i^2 de la formule de W_i. Deuxièmement, elle peut entraîner pendant les pas consécutifs une augmentation de la grandeur de sortie, ce qui entraîne également l'accroissement de la fonction de pertes. La commande optimale u_i^* résulte d'un compromis entre les tendances contradictoires exposées ci-dessus.

CHAPITRE VII

CERTAINES APPLICATIONS DE LA THÉORIE DE LA COMMANDE DUALE

§ 1. Synthèse d'un modèle auto-adaptatif optimal

La théorie des décisions statistiques permet de construire des méthodes optimales d'estimation des caractéristiques des systèmes de commande. D'après l'information qu'on possède au début de l'expérience, les problèmes qui se présentent dans le cadre général de la définition des caractéristiques d'un système peuvent être rangés en deux classes :

1. Définition de la structure inconnue du système.
2. Recherche de certains paramètres de l'opérateur connu du système.

Dans les conditions réelles, la définition des caractéristiques n'a lieu que si le système fonctionne normalement, que la fonction de commande principale ne soit pas compromise et, généralement, en présence d'un niveau de bruits élevé.

Le premier problème qui consiste à déterminer la structure inconnue d'un système en présence des perturbations peut être énoncé comme le problème d'estimation d'un opérateur de liaison des processus aléatoires. Ses méthodes de résolution se basent sur la théorie des fonctions aléatoires étudiée surtout par A Kolmogorov [1.14], N. Wiener [1.15].

La deuxième classe, plus restreinte, présente elle aussi un grand intérêt pratique. En effet, souvent le comportement d'un système peut être décrit, au moins en première approximation, en partant des considérations purement physiques. Dans ce cas l'expérience vise à rechercher les valeurs des paramètres justifiant l'hypothèse adoptée sur la structure du système.

Dans ce qui suit nous n'allons examiner que les problèmes de la deuxième classe. Deux approches différentes se font actuellement jour [7.1]. La première se ramène à définir analytiquement les caractéristiques à l'aide d'un calculateur, analogique ou numérique, la deuxième à adapter au système son modèle dynamique.

La figure 7.1 représente le schéma fonctionnel d'un système de définition analytique des caractéristiques. O désigne un objet caractérisé par le vecteur des paramètres $\bar{\mu}$; G et H sont les chaînes d'ac-

tion qui canalisent l'information sur les signaux d'entrée et de sortie vers le calculateur C. Dans le cas général, G et H sont soumises aux perturbations aléatoires $g(t)$ et $h(t)$. Le calculateur a pour tâche de trouver l'estimation $\bar{d}$ du vecteur des paramètres $\bar{\mu}$ sur la base de l'étude des processus aléatoires $y(t)$ et $z(t)$. Dans le cas de la définition analytique, le programme des calculs s'établit de façon à minimiser l'écart entre les caractéristiques calculées et les caractéristiques réelles.

Ainsi formulé, le problème s'apparente au problème de la réalisation optimale des signaux en présence des perturbations. La posi-

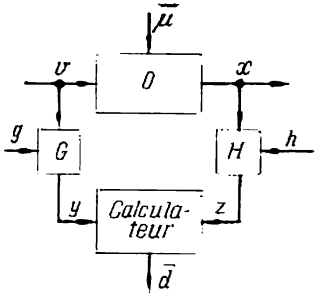


Fig. 7.1.

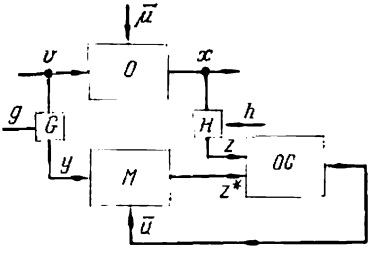


Fig. 7.2.

tion générale de ce problème et les diverses méthodes de résolution pour le temps discret et continu sont données dans [5.22, 1.10].

Si la théorie des communications examine d'habitude les cas où le même signal utile parcourt toutes les chaînes [7.2, 7.3], la définition analytique des caractéristiques suppose une différence essentielle de l'information transmise par des chaînes différentes. La solution générale de ce problème et des exemples sont donnés par [7.4 à 7.6].

Nous avons déjà dit que la deuxième méthode de définition des caractéristiques consiste à adapter au système son modèle dynamique : on construit un modèle réel, souvent plus simple que le système lui-même, qui est commandé de façon à minimiser une certaine fonction (ou une fonctionnelle) de l'écart entre les signaux de sortie du modèle et ceux du système. Les paramètres cherchés sont lus sur le modèle une fois son réglage achevé. Cette méthode est illustrée par la figure 7.2, où M est le modèle du système, OC l'organe de commande, G et H les chaînes d'action, $\bar{u}$ le vecteur des commandes.

En conservant les notations du chapitre précédent, examinons un système continu à états discrets. Supposons que l'opérateur du modèle soit choisi. Dans le cas général, la forme de l'opérateur du modèle et de celui du système peut être différente. Cela signifie qu'un systè-

me de structure définie est approché par un modèle de structure différente. Il se peut également que les dimensions des vecteurs $\bar{\mu}$ et $\bar{u}$ ne coïncident pas non plus : le réglage est donc appliqué à certains et non pas à tous les paramètres du modèle.

Le modèle est un système à paramètres variables et son signal de sortie z_s^* dépend à l'instant donné du réglage précédent.

Des suites des signaux aléatoires $\vec{z}_{s-1}^*$ et $\vec{z}_{s-1}$ sont appliquées aux entrées de l'organe de commande muni d'une mémoire qui stocke en outre toute la matrice des commandes $\vec{\bar{u}}_{s-1}$ que l'OC a élaborée. Le problème se pose d'obtenir la structure et la règle de décision $\Gamma_s(\bar{u}_s|\vec{\bar{u}}_{s-1}, \vec{z}_{s-1}^*, \vec{z}_{s-1})$ d'un organe de commande qui après l'étude des suites $\vec{z}_{s-1}^*$, $\vec{z}_{s-1}$ et de la matrice $\vec{\bar{u}}_{s-1}$ fournisse la commande

$$\bar{u}_s = \bar{u}_s(\vec{\bar{u}}_{s-1}, \vec{z}_{s-1}^*, \vec{z}_{s-1}), \tag{7.1}$$

qui soit optimale dans le sens du critère du minimum du risque total

$$R = M\left\{\sum_{s=1}^{n} W_s(s, z_s^*, x_s)\right\} = \sum_{s=1}^{n} M\{W_s\} = \sum_{s=1}^{n} R_s. \tag{7.2}$$

En principe on peut étudier également des fonctions de pertes élémentaires W_s plus complexes dépendant non pas des valeurs instantanées z_s^* et x_s des signaux de sortie du modèle et du système mais des vecteurs

$$\vec{z}_{sl}^* = (z_s^*, z_{s+1}^*, \ldots, z_{s+l}^*), \quad \vec{x}_{sl} = (x_s, x_{s+1}, \ldots, x_{s+l}).$$

Un organe de commande est dit optimal s'il forme la suite des commandes $\bar{u}_s^* = \bar{u}_s^*(\vec{\bar{u}}_{s-1}^*, \vec{z}_{s-1}^*, \vec{z}_{s-1})$ minimisant le risque total. Ainsi, la conduite du modèle consiste à minimiser l'écart entre les processus aléatoires étalon $\vec{x}_s$ et commandé $\vec{z}_s^*$ dans l'intervalle de temps $0 \leqslant s \leqslant n$.

Il faut pourtant noter que même en l'absence de perturbations dans la chaîne G, les processus aléatoires $\vec{x}_s$ et $\vec{z}_s^*$ peuvent être proches sans que cela soit vrai de tous les paramètres du modèle et du système, ce qui sans aucun doute constitue un inconvénient de la méthode faisant appel à un modèle auto-adaptatif. Il est pourtant fréquent que cette méthode est la seule possible. Il en est ainsi, par exemple, quand on a à approcher un système réel complexe par un modèle de structure plus simple.

Supposons que le signal $\vec{v}_s$ et les perturbations $\vec{g}_s$ et $\vec{h}_s$ forment des suites de variables aléatoires réparties suivant des lois fixes et

que les processus $\vec{v}_s$, $\vec{g}_s$, $\vec{h}_s$ ainsi que le vecteur des paramètres $\bar{\mu}$ sont indépendants les uns des autres. Les modes de combinaison du signal et du bruit dans les chaînes d'action G et H sont supposés connus et invariables, les chaînes elles-mêmes étant sans mémoire. Toutes les relations s'établissent pour le cas d'un système stationnaire défini par le vecteur des variables aléatoires

$$\bar{\mu}_s = \bar{\mu} = (\mu_1, \mu_2, \ldots, \mu_k).$$

La généralisation au cas des processus markoviens se fait sans peine de même que dans le chapitre précédent.

L'expression du risque élémentaire, i.e. du risque associé à la fonction de pertes $W_s(s, l, \vec{z}^*_{sl}\ \vec{x}_{sl})$, s'écrit

$$R_s = \int\limits_{\Omega(\vec{z}^*_{sl}, \vec{x}_{sl}, \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1})} W_s(s, l, \vec{z}^*_{sl}, \vec{x}_{sl}) \times$$
$$\times P(\vec{z}^*_{sl}, \vec{x}_{sl} \mid \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1})\, P(\vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1})\, d\Omega, \qquad (7.3)$$

où $\vec{z}^*_{sl} = (z^*_s, z^*_{s+1}, \ldots, z^*_{s+l})$ est le vecteur des signaux de sortie du modèle entre s et $(s+l)$ tandis que $\vec{x}_{sl} = (x_s, x_{s+1}, \ldots, x_{s+l})$ est le vecteur des signaux de sortie du système aux mêmes instants.

Avant de passer à la transformation des termes sous le signe somme de (7.3), considérons en bref les exigences imposées à la règle de décision $\Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1}, \vec{z}^*_{s-1}, \vec{z}_{s-1})$ de l'organe de commande.

Cette règle peut être régulière ou stochastique, mais son trait essentiel est de permettre la définition de la commande optimale $\bar{u}^*_s$ à l'aide des seules données d'observation $\vec{\bar{u}}_{s-1}$, $\vec{z}^*_{s-1}$, $\vec{z}_{s-1}$. Autrement dit, nous cherchons un organe de commande dans la classe des systèmes qui vérifient la relation

$$\Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1}, \vec{z}^*_{s-1}, \vec{z}_{s-1}, \bar{\mu}, \vec{v}_k) = \Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1}, \vec{z}^*_{s-1}, \vec{z}_{s-1}). \qquad (7.4)$$

La relation (7.4) est la condition de la réalisabilité physique du système.

Ecrivons le deuxième facteur de l'expression sous le signe somme (7.3) sous la forme suivante :

$$P(\vec{z}^*_{sl}, \vec{x}_{sl} \mid \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1}) =$$
$$= \int\limits_{\Omega(\bar{u}_s, \bar{\mu}, \vec{v}_{s+l})} P(\vec{z}^*_{sl}, \vec{x}_{sl}, \bar{u}_s, \bar{\mu}, \vec{v}_{s+l} \mid \vec{z}^*_{s-1}, \vec{z}_{s-1}\ \vec{\bar{u}}_{s-1})\, d\Omega. \qquad (7.5)$$

Ici $\vec{v}_{s+l} = (v_1, v_2, \ldots, v_s, v_{s+1}, \ldots, v_{s+l})$ est le vecteur des signaux d'entrée du système entre le premier et le $(s+l)$-ième instants.

D'après le théorème du produit des probabilités

$$P(\vec{z}^*_{sl}, \vec{x}_{sl}, \bar{u}_s, \bar{\mu}, \vec{v}_{s+l} \mid \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1}) =$$
$$= P(\bar{\mu}, \vec{v}_{s+l} \mid \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1})\, P(\vec{z}^*_{sl}, \vec{x}_{sl}, \bar{u}_s \mid \bar{\mu}, \vec{v}_{s+l}, \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1}). \tag{7.6}$$

Mettons le deuxième facteur de la formule (7.6) également sous la forme d'un produit

$$P(\vec{z}^*_{sl}, \vec{x}_{sl}, \bar{u}_s \mid \bar{\mu}, \vec{v}_{s+l}, \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1}) =$$
$$= P(\bar{u}_s \mid \bar{\mu}, \vec{v}_{s+l}, \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1})\, P(\vec{z}^*_{sl}, \vec{x}_{sl} \mid \vec{\bar{u}}_s, \bar{\mu}, \vec{v}_{s+l}, \vec{z}^*_{s-1}, \vec{z}_{s-1}). \tag{7.7}$$

Considérons la formule (7.7). Conformément à la condition (7.4) son premier facteur est la règle de décision $\Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1}, \vec{z}^*_{s-1}, \vec{z}_{s-1})$ de l'organe de commande. Ensuite, les perturbations dans les chaînes G et H étant statistiquement indépendantes, le deuxième facteur de la formule (7.7) peut s'écrire

$$P(\vec{z}^*_{sl}, \vec{x}_{sl} \mid \vec{\bar{u}}_s, \bar{\mu}, \vec{v}_{s+l}, \vec{z}^*_{s-1}, \vec{z}_{s-1}) =$$
$$= P(\vec{z}^*_{sl} \mid \vec{z}^*_{s-1}, \vec{\bar{u}}_s, \vec{v}_{s+l})\, P(\vec{x}_{sl} \mid \vec{z}_{s-1}, \bar{\mu}, \vec{v}_{s+l}). \tag{7.8}$$

Examinons le premier terme de la relation (7.8). Vu que $\vec{z}^*_{sl} = (z^*_s, z^*_{s+1}, \ldots, z^*_{s+l})$, il se met sous la forme suivante:

$$P(z^*_s, z^*_{s+1}, \ldots, z^*_{s+l} \mid z^*_{s-1}, \vec{v}_{s+l}, \vec{\bar{u}}_s) =$$
$$= \prod_{j=0}^{l} P(z^*_{s+j} \mid \vec{z}^*_{s+j-1}, \vec{\bar{u}}_s, \vec{v}_{s+j}). \tag{7.9}$$

Le deuxième facteur de (7.8) peut s'écrire de façon analogue. Mais la figure 7.2 montre qu'avec les vecteurs $\bar{\mu}$ et $\vec{v}_{s+j}$ fixés la densité de probabilité $P(x_{s+j} \mid \bar{\mu}, \vec{v}_{s+j}, \vec{z}_{s-1})$ se transforme en fonction delta

$$P(x_{s+j} \mid \bar{\mu}, \vec{v}_{s+j}, \vec{z}_{s-1}) = \delta(x_{s+j} - x^*_{s+j}), \tag{7.10}$$

où $x^*_{s+j} = x^*_{s+j}(\bar{\mu}, \vec{v}_{s+j})$ est la valeur du signal de sortie associée à l'ensemble fixé $(\bar{\mu}, \vec{v}_{s+j})$.

Ainsi, l'expression (7.7) devient

$$P(\vec{z}^*_{sl}, \vec{x}_{sl}, \bar{u}_s \mid \bar{\mu}, \vec{v}_{s+l}, \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1}) =$$
$$= \Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1}, \vec{z}^*_{s-1}, \vec{z}_{s-1}) \prod_{j=0}^{l} P(z^*_{s+j} \mid \vec{z}^*_{s+j-1}, \vec{\bar{u}}_s, \vec{v}_{s+j}) \times$$
$$\times \prod_{j=0}^{l} \delta(x_{s+j} - x^*_{s+j}), \quad 0 \leqslant j \leqslant l. \tag{7.11}$$

Revenons au premier facteur de la formule (7.6). Constatons d'abord que pour $i < k$ la condition de réalisabilité physique entraîne l'indépendance de z^*_i et z_i de $\vec{v}_k$ du fait que les signaux de sortie des systèmes réels ne peuvent dépendre des signaux d'entrée futurs.

D'après le théorème de Bayes

$$P(\bar{\mu}, \vec{v}_{s+l} \mid \vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1}) =$$
$$= P(\bar{\mu}, \vec{v}_{s+l}) \frac{P(\vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1} \mid \bar{\mu}, \vec{v}_{s+l})}{P(\vec{z}^*_{s-1}, \vec{z}_{s-1}, \vec{\bar{u}}_{s-1})}. \tag{7.12}$$

En récrivant le numérateur de (7.12) tout comme on l'a fait dans le chapitre VI (cf. (6.28)) et en effectuant les substitutions successives, on obtient l'expression suivante du risque élémentaire R_s:

$$R_s = \int\limits_{\Omega(\vec{z}^*_{s+l}, \vec{z}_{s-1}, \vec{\bar{u}}_s, \bar{\mu}, \vec{x}_{s+l})} W_s(s, l, \vec{z}^*_{sl}, \vec{x}_{sl}) P(\bar{\mu}) P(\vec{v}_{s+l}) \times$$
$$\times \prod_{i=1}^{s-1} P(z_i \mid \bar{\mu}, \vec{v}_i, \vec{z}_{i-1}) \prod_{i=1}^{s-1} P(z^*_i \mid \vec{\bar{u}}_i, \vec{v}_i, \vec{z}^*_{i-1}) \times$$
$$\times \prod_{j=0}^{l} P(z^*_{s+j} \mid \vec{\bar{u}}_s, \vec{v}_{s+j}, \vec{z}^*_{s+j-1}) \prod_{i=2}^{s} \Gamma_i(\bar{u}_i \mid \vec{\bar{u}}_{i-1}, \vec{z}^*_{i-1}, \vec{z}_{i-1}) \Gamma_1(\bar{u}_1 \mid \bar{z}^*_0)\, d\Omega, \tag{7.13}$$

où

$$\vec{z}^*_{s+l} = (\bar{z}^*_0, z^*_1, z^*_2, \ldots, z^*_{s-1}, z^*_s, z^*_{s+1}, \ldots, z^*_{s+l}) =$$
$$= (\vec{z}^*_{s-1}, \vec{z}^*_{s+l}), \quad 1 \leqslant s \leqslant n,$$

et $\bar{z}^*_0$ sont les conditions initiales du modèle.

Le risque total en n pas est égal à la somme des risques élémentaires

$$R=\sum_{s=1}^{n} R_s=\sum_{s=1}^{n} \int\limits_{\Omega(\vec{z}^*_{s+l},\, \vec{z}_{s-1},\, \vec{\bar{u}}_s,\, \bar{\mu},\, \vec{v}_{s+l})} W_s(s,\, l,\, \vec{z}^*_{sl},\, \vec{x}_{sl}) \times$$

$$\times P(\bar{\mu})\, P(\vec{v}_{s+l}) \prod_{i=1}^{s-1} P(z_i \mid \vec{z}_{i-1},\, \bar{\mu},\, \vec{v}_i) \prod_{i=1}^{s-1} P(z_i^* \mid \vec{\bar{u}}_i,\, \vec{v}_i,\, \vec{z}^*_{i-1}) \times$$

$$\times \prod_{j=0}^{l} P(z^*_{j+s} \mid \vec{\bar{u}}_s,\, \vec{v}_{s+j},\, \vec{z}^*_{s+j-1}) \prod_{i=2}^{s} \Gamma_i(\bar{u}_i \mid \vec{\bar{u}}_{i-1},\, \vec{z}^*_{i-1},\, \vec{z}_{i-1}) \times$$

$$\times \Gamma_1(\bar{u}_1 \mid z_0^*)\, d\Omega. \qquad (7.14)$$

Choisissons les fonctions $\Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1},\, \vec{z}^*_{s-1},\, \vec{z}_{s-1})$ pour $2 \leqslant s \leqslant n$ sous la condition de minimum du risque total. Dans le cas général $\Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1},\, \vec{z}^*_{s-1},\, \vec{z}_{s-1})$ étant les densités de probabilité, elles doivent donc vérifier la condition de normalisation

$$\int\limits_{\Omega(\bar{u}_s)} \Gamma_s(\bar{u}_s \mid \vec{\bar{u}}_{s-1},\, \vec{z}^*_{s-1},\, \vec{z}_{s-1})\, d\Omega = 1.$$

Choisissons tout d'abord la dernière fonction $\Gamma_n(\bar{u}_n \mid \vec{\bar{u}}_{n-1},\, \vec{z}^*_{n-1},\, \vec{z}_{n-1})$. Celle-ci est déterminée de la condition de minimum du dernier terme R_n car seul R_n dépend de Γ_n. Supposons que toutes les autres fonctions $\Gamma_1, \Gamma_2, \ldots, \Gamma_{n-1}$, de même que les commandes respectives $\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_{n-1}$, sont trouvées. Ecrivons l'expression de R_n

$$R_n=\int\limits_{\Omega(\vec{z}^*_{n+l},\, \vec{z}_{n-1},\, \vec{\bar{u}}_n,\, \bar{\mu},\, \vec{v}_{n+l})} W_n(n,\, l,\, \vec{z}^*_{nl},\, \vec{x}_{nl})\, P(\bar{\mu})\, P(\vec{v}_{n+l}) \times$$

$$\times \prod_{i=1}^{n-1} P(z_i \mid \vec{z}_{i-1},\, \bar{\mu},\, \vec{v}_i) \prod_{i=1}^{n-1} P(z_i^* \mid \vec{\bar{u}}_i,\, \vec{v}_i,\, \vec{z}^*_{i-1}) \times$$

$$\times \prod_{j=0}^{l} P(z^*_{n+j} \mid \vec{\bar{u}}_n,\, \vec{v}_{n+j},\, \vec{z}^*_{n+j-1}) \times$$

$$\times \prod_{i=2}^{n} \Gamma_i(\bar{u}_i \mid \vec{\bar{u}}_{i-1},\, \vec{z}^*_{i-1},\, \vec{z}_{i-1})\, \Gamma_i(\bar{u}_1 \mid z_0^*)\, d\Omega. \qquad (7.15)$$

Introduisons les fonctions

$$\alpha_s = \int\limits_{\Omega(\vec{z}^*_{sl},\, \bar{\mu},\, \vec{v}_{s+l})} W_s(s,\, l,\, \vec{z}^*_{sl},\, \vec{x}_{sl})\, P(\bar{\mu})\, P(\vec{v}_{s+l}) \times$$

$$\times \prod_{i=1}^{s-1} P(z_i \mid \bar{\mu},\, \vec{v}_i,\, \vec{z}_{i-1}) \prod_{i=1}^{s-1} P(z^*_i \mid \vec{\bar{u}}_i,\, \vec{v}_i,\, \vec{z}^*_{i-1}) \times$$

$$\times \prod_{j=0}^{l} P(z^*_{s+j} \mid \vec{\bar{u}}_s,\, \vec{v}_{s+j},\, \vec{z}^*_{s+j-1})\, d\Omega \qquad (7.16)$$

et

$$\beta_s = \prod_{i=1}^{s} \Gamma_i. \qquad (7.17)$$

Compte tenu de ces notations avec $s = n$, on a

$$R_n \int\limits_{\Omega(\bar{u}_n,\, \vec{\bar{u}}_{n-1},\, \vec{z}^*_{n-1},\, \vec{z}_{n-1})} \alpha_n \beta_{n-1} \Gamma_n\, d\Omega = \int\limits_{\Omega(\vec{\bar{u}}_{n-1},\, \vec{z}^*_{n-1},\, \vec{z}_{n-1})} \beta_{n-1} \varkappa_n\, d\Omega, \qquad (7.18)$$

où

$$\varkappa_n = \int\limits_{\Omega(\bar{u}_n)} \alpha_n \Gamma_n (\bar{u}_n \mid \vec{\bar{u}}_{n-1},\, \vec{z}^*_{n-1},\, \vec{z}_{n-1})\, d\Omega. \qquad (7.19)$$

(7.18) montre que, β_{n-1} étant donnée (cf. 7.17), le minimum de R_n s'obtient pour une commande $\bar{u}_n$ qui minimise la fonction $\varkappa_n$c

Adoptons l'hypothèse suivante qui s'accorde parfaitement ave. l'interprétation physique du problème. Recherchons l'organe de commande dans la classe des systèmes munis de stratégies pures, c'est-à-dire des systèmes dont la règle de décision est une fonction δ. Dans ce cas

$$\Gamma_s (\bar{u}_s \mid \vec{\bar{u}}_{s-1},\, \vec{z}^*_{s-1},\, \vec{z}_{s-1}) = \delta (\bar{u}_s - \bar{u}^*_s), \quad 1 \leqslant s \leqslant n, \qquad (7.20)$$

où $\bar{u}^*_s$ est un vecteur des commandes optimales. En posant $s = n$, en portant (7.20) dans (7.19) et en intégrant par rapport à $\bar{u}_n$ on tombe sur

$$\varkappa^*_n = \alpha^*_n, \qquad (7.21)$$

où α^*_n est une fonction associée à la commande optimale $\bar{u}^*_n$. On en déduit immédiatement la règle de recherche du vecteur optimal $\bar{u}^*_n$ qui minimise le risque élémentaire R_n : le choix de ce vecteur est déterminé en minimisant la fonction $\alpha_n (\bar{u}_n,\, \vec{\bar{u}}_{n-1},\, \vec{z}^*_{n-1},\, \vec{z}_{n-1})$ par rapport à $\bar{u}_n$.

De plus il est clair que

$$\bar{u}_n^* = \bar{u}_n^* (\vec{\bar{u}}_{n-1}, \vec{z}_{n-1}^*, \vec{z}_{n-1}). \tag{7.22}$$

Des raisonnements analogues à ceux du § 2 du chapitre VI montrent qu'on obtient la commande optimale $\bar{u}_{n-k}^*$ en minimisant la fonction

$$\gamma_{n-k} = \alpha_{n-k} + \int\limits_{\Omega(z_{n-k}^*,\, z_{n-k})} \gamma_{n-k+1}^* \, d\Omega, \tag{7.23}$$

$\gamma_n^* = \alpha_n^*$, par rapport à $\bar{u}_{n-k}$. Dans ces conditions

$$\Gamma_{n-k} = \delta (\bar{u}_{n-k} - \bar{u}_{n-k}^*). \tag{7.24}$$

On calcule ainsi toutes les fonctions $\Gamma_n, \Gamma_{n-1}, \ldots, \Gamma_2$ sauf Γ_1 et toutes les commandes respectives $\bar{u}_n, \bar{u}_{n-1}, \ldots, \bar{u}_2$ sauf $\bar{u}_1$. Pour ce qui est du choix de la commande initiale $\bar{u}_1$ (et de la fonction Γ_1), deux cas peuvent se présenter :

1) la commande $\bar{u}_1$ est donnée et, par suite, son choix est inutile ;

2) le choix de la commande $\bar{u}_1$ se fait selon les règles générales.

Il est clair que la commande initiale $\bar{u}_1$ dépend des seuls renseignements a priori sur le système et non pas des données d'observation absentes à l'instant initial.

Considérons un cas particulier d'un grand intérêt lorsque les perturbations sont absentes dans la chaîne G et le signal d'entrée $\vec{v}$ est un processus aléatoire muni d'une loi de répartition connue. Dans ce cas, évidemment,

$$\prod_{i=1}^{s-1} P (z_i^* \,|\, \vec{\bar{u}}_i, \vec{v}_i, \vec{z}_{i-1}^*) = \prod_{i=1}^{s-1} \delta (z_i^* - \tilde{z}_i^*), \tag{7.25}$$

$$\prod_{j=0}^{l} P (z_{s+j}^* \,|\, \vec{\bar{u}}_s, \vec{v}_{s+j}, \vec{z}_{s+j-1}^*) = \prod_{j=0}^{l} \delta (z_{s+j}^* - \tilde{z}_{s+j}^*), \tag{7.26}$$

où $\tilde{z}_i^* = \tilde{z}_i^* (\vec{\bar{u}}_i, \vec{v}_i, \bar{z}_0^*)$, $\tilde{z}_{s+j}^* = \tilde{z}_{s+j}^* (\vec{\bar{u}}_s, \vec{v}_{s+j}, \bar{z}_0^*)$ sont les valeurs réelles du signal de sortie du modèle aux instants i $(1 \leqslant i \leqslant s-1)$ et $s+j$ $(0 \leqslant j \leqslant l)$ respectivement.

En portant (7.25) et (7.26) dans l'expression (7.13) du risque élémentaire et en intégrant par rapport à $\vec{z}_{s+l}^*$, on a

$$R_s = \int\limits_{\Omega(\vec{z}_{s-1},\, \vec{\bar{u}}_s,\, \bar{\mu},\, \vec{v}_{s+l})} W_s (s, l, \vec{v}_{s+l}, \bar{\mu}, \vec{\bar{u}}_s, \bar{z}_0^*) \times$$

$$\times P (\bar{\mu}) P (\vec{v}_{s+l}) \prod_{i=1}^{s-1} P (z_i \,|\, \bar{\mu}, \vec{v}_i, \vec{z}_{i-1}) \times$$

$$\times \prod_{i=2}^{s} \Gamma_i (\bar{u}_i \,|\, \vec{\bar{u}}_{i-1}, \vec{z}_{i-1}, \vec{v}_{i-1}, \bar{z}_0^*) \, \Gamma_1 (\bar{u}_1 \,|\, \bar{z}_0^*) \, d\Omega. \tag{7.27}$$

De même que dans le cas précédent, le choix des commandes optimales $\bar{u}_n^*$, $\bar{u}_{n-1}^*$, ..., $\bar{u}_1^*$ se fait par itération.

Introduisons la fonction

$$\alpha_s = \int\limits_{\Omega(\bar{\mu},\, \vec{v}_{s+l})} W_s(s,\, l,\, \vec{v}_{s+l},\, \bar{\mu},\, \vec{\bar{u}}_s,\, \bar{z}_0^*)\, P(\bar{\mu})\, P(\vec{v}_{s+l}) \times$$

$$\times \prod_{i=1}^{s-1} P(z_i \mid \bar{\mu},\, \vec{v}_i,\, \vec{z}_{i-1})\, d\Omega, \qquad (7.28)$$

où

$$v_{s+l} = (v_1,\, v_2,\, \ldots,\, v_{s-1},\, v_s,\, v_{s+1},\, \ldots,\, v_{s+l}) = (\vec{v}_{s-1},\, \vec{v}_{sl}). \qquad (7.29)$$

Un raisonnement identique à celui du cas précédent montre que la dernière commande optimale $\bar{u}_n^*$ se détermine en minimisant α_n de (7.28) par rapport à $\bar{u}_n$. Dans ces conditions

$$\bar{u}_n^* = \bar{u}_n^*(\vec{\bar{u}}_{n-1},\, \vec{z}_{n-1},\, \vec{v}_{n-1},\, \bar{z}_0^*) \qquad (7.30)$$

et

$$\Gamma_n(\bar{u}_n \mid \vec{\bar{u}}_{n-1},\, \vec{z}_{n-1},\, \vec{v}_{n-1},\, \bar{z}_0^*) = \delta(\bar{u}_n - \bar{u}_n^*). \qquad (7.31)$$

On établit de même que pour obtenir la commande optimale $\bar{u}_{n-k}^*$ il faut minimiser la fonction

$$\gamma_{n-k} = \alpha_{n-k} + \int\limits_{\Omega(z_{n-k},\, v_{n-k})} \gamma_{n-k+1}^*\, d\Omega, \qquad (7.32)$$

par rapport à $\bar{u}_{n-k}$, où $n-k > 1$. De plus

$$\alpha_n^* = \gamma_n^*,$$

$$\Gamma_{n-k} = \delta(\bar{u}_{n-k} - \bar{u}_{n-k}^*).$$

Les formules deviennent encore plus simples si le signal d'entrée $\vec{v}$ du modèle et du système est un processus déterministe variant de façon connue en fonction du temps. Ce cas où une perturbation non contrôlable agit seulement à la sortie du système tandis qu'aux entrées du modèle et du système est appliqué un même signal déterministe (discret ou continu) est le mieux étudié dans la littérature [7.1].

Il est clair que dans ce cas

$$P(\vec{v}_{s+l}) = \prod_{i=1}^{s+l} \delta(v_i - v_i^*), \qquad (7.33)$$

où v_i^* est la valeur réelle du signal d'entrée à l'instant $t = i$.

Revenons à la condition (7.4) à laquelle doit satisfaire la règle de décision Γ_s de l'organe de commande optimal OC. Nous avons précisé plus haut que le sens physique de cette condition est dans ce que la commande $\bar{u}_s$, élaborée dans l'organe de commande à l'instant donné, n'est définie que par l'information fournie aux instants antérieurs.

Pourtant, si le signal d'entrée $\vec{v}$ est un processus déterministe dont on connaît la loi de variation dans le temps, l'organe de commande possède une information plus complète (par rapport à l'état précédent) sur le système.

En effet, l'organe de commande dispose alors à tout instant d'une information non seulement sur le comportement antérieur du système, mais encore prévoit les valeurs futures du processus à l'entrée du système, et donc peut assurer un choix plus heureux de la commande optimale $\bar{u}_s^*$.

Dans ces conditions l'égalité (7.4) devient

$$\Gamma_s(u_s \mid \vec{\bar{u}}_{s-1}, \vec{z}^*_{s-1}, \vec{z}_{s-1}, \bar{\mu}, \vec{v}^*_k) = \Gamma_s(u_s \mid \vec{\bar{u}}_{s-1}, \vec{z}_{s-1}, \vec{v}^*_k, \bar{z}^*_0), \qquad (7.34)$$

où $k \geqslant s$. Par conséquent il se peut que la commande $\bar{u}_s$ dépende des valeurs futures du signal d'entrée. Ce phénomène tient physiquement à ce que dans le choix de la commande optimale le système tend à utiliser toute l'information fournie, y compris celle sur le comportement futur du signal d'entrée déterministe.

En portant (7.33) dans (7.27), en intégrant par rapport à $\vec{v}_{s+l}$ et compte tenu de la condition (7.34), on aboutit à

$$R_s = \int\limits_{\Omega(\vec{z}_{s-1}, \vec{\bar{u}}_s, \bar{\mu})} W_s(s, l, \vec{v}^*_{s+l}, \bar{\mu}, \vec{\bar{u}}_s, \bar{z}^*_0)\, P(\bar{\mu}) \prod_{i=1}^{s-1} P(z_i \mid \bar{\mu}, \vec{v}_i, \vec{z}_{i-1}) \times$$
$$\times \prod_{i=2}^{s} \Gamma_i(\bar{u}_i \mid \vec{\bar{u}}_{i-1}, \vec{z}_{i-1}, \vec{v}^*_{i+p}, \bar{z}^*_0)\, \Gamma_1(\bar{u}_1 \mid \bar{z}^*_0, \vec{v}^*_{1+p})\, d\Omega. \qquad (7.35)$$

Dans le cas considéré la recherche des commandes optimales $\bar{u}_n^*, \bar{u}_{n-1}^*, \ldots, \bar{u}_1^*$ se fait toujours par itération.

Introduisons la fonction

$$\alpha_s = \int\limits_{\Omega(\bar{\mu})} W_s(s, l, \vec{v}^*_{s+l}, \bar{\mu}, \vec{\bar{u}}_s, z^*_0)\, P(\mu) \times$$
$$\times \prod_{i=1}^{s-1} P(z_i \mid \bar{\mu}, \vec{v}_i, \vec{z}_{i-1})\, d\Omega. \qquad (7.36)$$

On montre aisément que la commande optimale $\bar{u}_n^*$ est fournie par la minimisation de α_n par rapport à $\bar{u}_n$. De plus

$$\bar{u}_n^* = \bar{u}_n^* (\vec{u}_{n-1}, \vec{z}_{n-1}, \vec{v}_{n+l}^*, \bar{z}_0^*)$$

et

$$\Gamma_n = \delta (\bar{u}_n - \bar{u}_n^*).$$

Il n'est pas difficile de montrer que la commande optimale $\bar{u}_{n-k}^*$ s'obtient en minimisant par rapport à $\bar{u}_{n-k}$ la fonction

$$\gamma_{n-k} = \alpha_{n-k} + \int\limits_{\Omega(z_{n-k})} \gamma_{n-k+1}^* \, d\Omega,$$
$$n > n-k > 1, \qquad (7.37)$$

où $\gamma_n^* = \alpha_n^*$, $\Gamma_{n-k} = \delta(\bar{u}_{n-k} - \bar{u}_{n-k}^*)$.

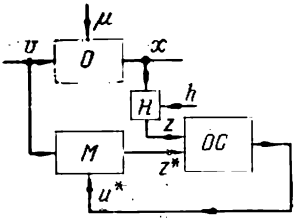


Fig. 7.3.

Considérons à titre d'exemple la synthèse d'un modèle auto-adaptatif de l'intégrateur. Supposons que la chaîne G ne donne pas lieu à une perturbation et que les signaux d'entrée du modèle et du système constituent le même processus aléatoire $\vec{v}$ (fig. 7.3) qui est une suite des échantillons indépendants de loi de répartition invariable

$$P(v_s) = \frac{1}{\sigma_v \sqrt{2\pi}} \exp\left\{-\frac{(v_s - m_v)^2}{2\sigma_v^2}\right\}. \qquad (7.38)$$

D'après la figure 7.3 *) on peut écrire pour le s-ième pas

$$z_s = x_s + h_s = \mu \sum_{i=1}^{s} v_i + x_0 + h_s, \qquad (7.39)$$

où μ est le gain aléatoire et x_0 la valeur initiale du signal de sortie de l'intégrateur.

En général, la condition initiale x_0, de même que le gain μ changent aléatoirement d'un processus à l'autre et on considérera donc que le signal de sortie dépend non pas de la seule variable aléatoire μ mais du vecteur aléatoire (μ, v_0). Toitefois, dans l'exemple considéré, admettons que l'état initial du système est connu avec précision. Cette hypothèse nous sera utile dans ce qui suit.

Supposons que la perturbation $\vec{h}$ soit une suite des variables aléatoires indépendantes de densité de probabilité invariable

$$q(h_s) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{h_s^2}{2\sigma_h^2}\right\}, \qquad (7.40)$$

*) Un schéma analogue permettant la définition analytique des paramètres sera appelé «ouvert».

et que le paramètre aléatoire μ vérifie la loi normale

$$q_1(\mu) = \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2}\right\}. \tag{7.41}$$

Le gain du modèle change d'un pas à l'autre. Compte tenu de tout ce qui a été dit, l'équation du modèle peut se mettre sous la forme

$$z_s^* = u_s v_s + z_{s-1}^*. \tag{7.42}$$

Choisissons la fonction de pertes quadratique

$$W_s = (z_s^* - x_s)^2 = (u_s v_s + z_{s-1}^* - \mu \sum_{j=1}^{s} v_j - x_0)^2. \tag{7.43}$$

Posons $s = n$ et écrivons la fonction α_n à minimiser par rapport à u :

$$\alpha_n = \frac{1}{\sigma_\mu \sigma_v \sigma_h^{n-1} (2\pi)^{0,5(n+1)}} P(\vec{v}_{n-1}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Big(u_n v_n + z_{n-1}^* - \\ - \mu \sum_{j=1}^{n} v_j - x_0\Big)^2 \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma_\mu^2} - \frac{(v_n - m_v)^2}{2\sigma_v^2} - \right. \\ \left. - \frac{1}{2\sigma_h^2} \sum_{i=1}^{n-1} \Big(z_i - \mu \sum_{j=1}^{i} v_i - x_0\Big)^2\right\} d\mu\, dv_n. \tag{7.44}$$

Des calculs élémentaires mais très longs amènent

$$u_n^* = d_{n-1}^* + \frac{m_v}{\sigma_v^2 + m_v^2} (\rho_{n-1} d_{n-1}^* - z_{n-1}^* + x_0), \tag{7.45}$$

où

$$\left.\begin{aligned} d_{n-1}^* &= \frac{\sigma_h^2 \mu_0 + \sigma_\mu^2 \sum_{i=1}^{s} \theta_i \rho_i}{\sigma_h^2 + \sigma_\mu^2 \sum_{i=1}^{s} \rho_i^2}, \\ \rho_i &= \sum_{j=1}^{i} v_j, \\ \theta_i &= z_i - x_0. \end{aligned}\right\} \tag{7.46}$$

Ici d_{n-1}^* est l'estimation optimale établie par la méthode analytique après l'observation de $\vec{v}_{n-1}$, $\vec{z}_{n-1}$ [7.7, 7.8]. Malheureusement seule la commande optimale terminale u_n^* peut s'obtenir sous forme explicite.

Si l'on suppose que le signal d'entrée est un processus déterministe variant de façon continue, la situation devient bien meilleure. Dans ce cas, $m_v = v_n$, $\sigma_v^2 = 0$ et la formule (7.45) devient

$$u_n^* = d_{n-1}^* + \frac{1}{v_n}(\rho_{n-1}d_{n-1}^* - z_{n-1}^* + x_0). \tag{7.47}$$

Le schéma comportant le modèle est en l'occurrence neutre et, par conséquent, la commande optimale u^* pour n'importe quel $s > 1$ se calcule d'après une formule analogue à (7.47), c'est-à-dire

$$u_s^* = d_{s-1}^* + \frac{1}{v_s}(\rho_{s-1}d_{s-1}^* - z_{s-1}^* + x_0). \tag{7.48}$$

La commande optimale initiale u_1^* choisie avant l'observation s'écrit

$$u_1^* = \mu_0 - \frac{z_0^* - x_0}{v_1}, \tag{7.49}$$

où z_0^* est la valeur initiale du signal de sortie du modèle et μ_0 l'espérance mathématique de la densité de probabilité du paramètre μ.

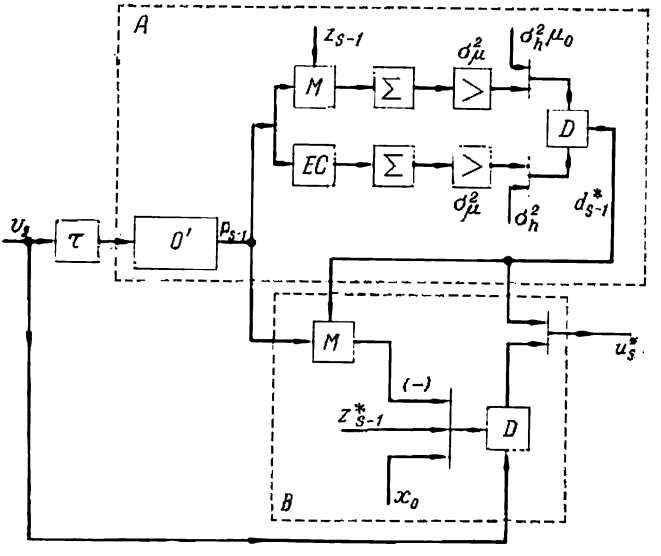


Fig. 7.4.

Dans ce qui suit nous allons étudier le système au signal d'entrée déterministe.

La figure 7.4 représente le schéma fonctionnel d'un organe de commande optimal construit d'après les formules (7.48) et (7.49).

L'organe de commande est composé de deux blocs : bloc A de recherche de l'estimation optimale d^*_{s-1} et bloc B de définition de la commande optimale u^*_s. O' désigne ici l'intégrateur à gain unitaire ; Σ les additionneurs cumulatifs ; M les multiplieurs ; D les diviseurs ; EC l'élévateur au carré ; τ le circuit de retard d'un pas.

Les formules (7.48) et (7.49) montrent que pour $v_s \to 0$ la commande optimale $u^*_s \to \infty$. Pour obtenir un système physiquement réalisable il faut introduire la contrainte $|u^*_s| \leqslant Q$ et comparer chaque fois la commande obtenue avec le seuil Q.

Pour simplifier l'analyse ultérieure supposons le nombre Q et le signal d'entrée si grands que la condition $|u^*_s| \leqslant Q$ soit vérifiée pour toute commande donnée par la formule (7.48). La commande optimale u^*_s peut s'écrire alors sous une forme un peu différente. Il est clair que

$$z^*_{s-1} = v_{s-1}u_{s-1} + z^*_{s-2}. \tag{7.50}$$

Soit $u_{s-1} = u^*_{s-1}$. D'après (7.48) nous avons alors pour z_{s-1}

$$z^*_{s-1} = v_{s-1}u^*_{s-1} + z^*_{s-2} = v_{s-1}\left[d^*_{s-2} + \frac{1}{v_{s-1}}(\rho_{s-2}d^*_{s-2} - z^*_{s-2} + x_0)\right] + \\ + z^*_{s-2} = v_{s-1}d^*_{s-2} + \rho_{s-2}d^*_{s-2} + x_0. \tag{7.51}$$

En portant (7.51) dans (7.48) il vient

$$u^*_s = d^*_{s-1} + \frac{\rho_{s-1}}{v_s}(d^*_{s-1} - d^*_{s-2}), \quad s = 2, 3, \ldots \tag{7.52}$$

Considérons les formules (7.46) et (7.52). Les fonctions u^*_s et d^*_s sont les estimations du même gain aléatoire μ, mais établies en fonction des objectifs de commande différents : u^*_s se calcule à partir de la condition de minimum de l'espérance mathématique de la somme des carrés des écarts entre les signaux de sortie du modèle et ceux du système, alors que d^*_s s'établit à partir de la condition de minimum de la variance moyenne de l'estimation. Essayons de comparer ces estimations. Notons tout d'abord que d^*_{s-1} se définit après l'obtention du vecteur $(\vec{v}_{s-1}, \vec{z}_{s-1})$, alors que pour prédire u^*_s on se base sur l'ensemble $(\vec{v}_s, \vec{z}_{s-1})$. Le processus discret $\vec{v}$ étant déterministe, les estimations u^*_s et d^*_{s-1} se calculent pour le même volume d'information sur le système et, par conséquent, ce sont ces estimations qu'il faut comparer.

Considérons la convergence des u^*_{s+1} et d^*_s vers μ. Soit $\bar{\mu}$ la valeur vraie du paramètre μ dans le cycle de régulation considéré. Les estimations u^*_{s+1} et d^*_s étant des fonctions de réseau aléatoires, nous examinons ci-dessous la convergence vers μ non pas des esti-

mations elles-mêmes, mais de leurs espérances mathématiques conditionnelles $M\{u^*_{s+1} \mid \bar{\mu}\}$ et $M\{d^*_s \mid \bar{\mu}\}$.

Comme il s'ensuit des formules (7.46) et (7.52), d^*_s et u^*_{s+1} dépendent non pas des variances σ^2_μ et σ^2_h mais de leur rapport $\beta^2 = \sigma^2_h/\sigma^2_\mu$. La figure 7.5 représente la variation de $M\{u^*_{s+1} \mid \bar{\mu}\}$ (courbe *1*) et de $M\{d^*_s \mid \bar{\mu}\}$ (courbe *2*) en fonction du nombre d'observations, pour les valeurs du paramètre $\beta^2 = 0{,}1$; 1 ; 10 avec $\mu_0 = x_0 = z^*_0 = 0$ et le saut unitaire à l'entrée du système : $v_1 = v_2 = \ldots = 1$. Il faut prendre en considération que toutes ces courbes sont des enveloppes des fonctions de réseau définies seulement à des instants exprimés par des nombres entiers.

La figure 7.5 montre que pour des μ_0, x_0, z^*_0 données, les fonctions $M\{u^*_{s+1} \mid \bar{\mu}\}$ et $M\{d^*_s \mid \bar{\mu}\}$ convergent vers la valeur vraie $\bar{\mu}$ de façon à peu près identique mais suivant des lois différentes. Dans le cas d'un modèle auto-adaptatif on est en présence d'une surrégulation, alors que la convergence de $M\{d^*_s \mid \bar{\mu}\}$ vers $\bar{\mu}$ est monotone.

Le schéma à modèle auto-adaptatif possède une propriété curieuse. Ecrivons successivement quelques commandes optimales (formules (7.49) et (7.52)) :

$$\left.\begin{aligned} u^*_1 &= d^*_0 - \frac{z^*_0 - x_0}{\rho_1}\,. \\ u^*_2 &= \frac{\rho_2}{v_2}\, d^*_1 - \frac{\rho_1}{v_2}\, d^*_0, \\ &\ldots\ldots\ldots\ldots\ldots \\ u^*_s &= \frac{\rho_s}{v_s}\, d^*_{s-1} - \frac{\rho_{s-1}}{v_s}\, d^*_{s-2}. \end{aligned}\right\} \tag{7.53}$$

En multipliant chaque commande u^*_i par v_i et en additionnant terme à terme les deux membres des relations (7.53), on trouve le signal de sortie du modèle à la fin du s-ième pas

$$z^*_s = u^*_1 v_1 + u^*_2 v_2 + \ldots + u^*_s v_s = \rho_s d^*_{s-1} + x_0 - z^*_0. \tag{7.54}$$

D'après (7.54) le modèle n'accumule pas les erreurs de définition du gain, en sélectionnant chaque fois l'estimation la plus précise dans les conditions considérées, à savoir d^*_{s-1}, et en compensant l'influence des estimations précédentes d^*_0, d^*_1, . . ., d^*_{s-2} sur le signal de sortie du modèle.

Plus haut nous avons supposé connues les conditions initiales x_0 du système. Supposons maintenant qu'en les définissant nous avons commis une erreur Δ_x. Cette erreur entraîne des erreurs sur la définition des Δd^*_{s-1} et Δu^*_s.

Voyons comment Δd^*_{s-1} et Δu^*_s dépendent de Δ_x. Soit toujours $v_1 = v_2 = \ldots = 1$. La formule (7.46) permet de tirer

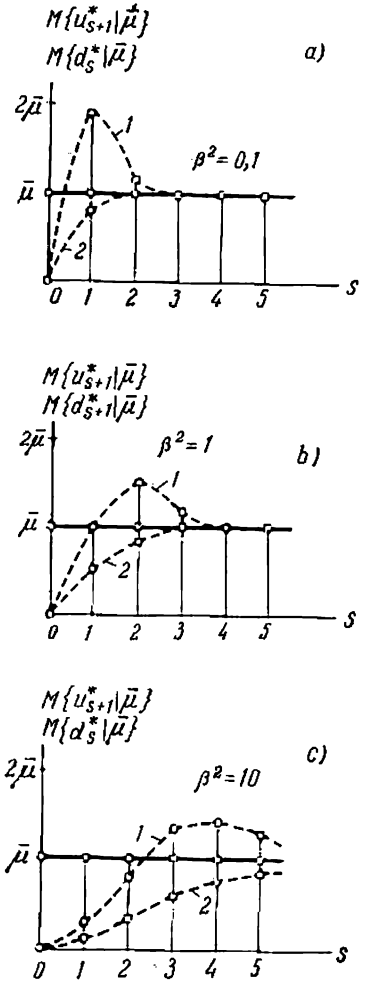


Fig. 7.5.

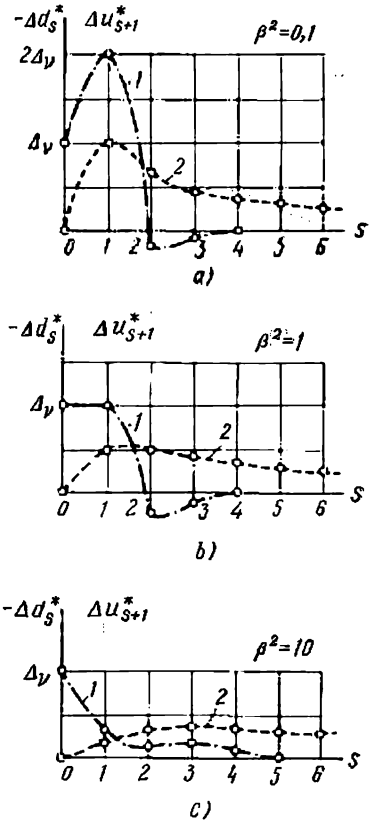


Fig. 7.6.

$$d^*_{s-1} = \frac{\beta^2\mu_0 + \sum\limits_{i=1}^{s-1} \rho_i z_i - x_0 \sum\limits_{i=1}^{s-1} \rho_i}{\beta^2 + \sum\limits_{i=1}^{s-1} \rho_i^2}. \qquad (7.55)$$

Avec $x = x_0 + \Delta_x$ l'erreur sur la définition de d^*_{s-1} s'écrit

$$\Delta d^*_{s-1} = -\Delta_x \frac{\sum\limits_{i=1}^{s-1} \rho_i}{\beta^2 + \sum\limits_{i=2}^{s-1} \rho_i^2}. \tag{7.56}$$

D'après (7.52) nous avons pour le schéma à modèle auto-adaptatif

$$\Delta u^*_s = -\Delta_x \left\{ \frac{\sum\limits_{i=1}^{s-1} \rho_i}{\beta^2 + \sum\limits_{i=1}^{s-1} \rho_i^2} + (s-1) \left[\frac{\sum\limits_{i=1}^{s-1} \rho_i}{\beta^2 + \sum\limits_{i=1}^{s-1} \rho_i^2} - \frac{\sum\limits_{i=1}^{s-2} \rho_i}{\beta^2 + \sum\limits_{i=1}^{s-2} \rho_i^2} \right] \right\}. \tag{7.57}$$

Considérons l'expression entre crochets. Sous les hypothèses faites ($v_i = 1$), on a

$$\sum_{i=1}^{s-1} \rho_i = \rho_1 + \rho_2 + \ldots + \rho_{s-1} = 1 + 2 + \ldots + (s-1) = \frac{s(s-1)}{2},$$

$$\sum_{i=1}^{s-1} \rho_i^2 = 1^2 + 2^2 + \ldots + (s-1)^2 = \frac{s(s-1)(2s-1)}{6}. \tag{7.58}$$

En substituant (7.58) dans (7.57), on obtient

$$\frac{\sum\limits_{i=1}^{s-1} \rho_i}{\beta^2 + \sum\limits_{i=1}^{s-1} \rho_i^2} - \frac{\sum\limits_{i=1}^{s-2} \rho_i}{\beta^2 + \sum\limits_{i=1}^{s-2} \rho_i^2} =$$

$$= \frac{1}{2} \left[\frac{s(s-1)}{\beta^2 + \frac{1}{6} s(s-1)(2s-1)} - \frac{(s-1)(s-2)}{\beta^2 + \frac{1}{6}(s-1)(s-2)(2s-3)} \right]. \tag{7.59}$$

A partir d'un certain s, la condition

$$\frac{(s-1)(s-2)(2s-3)}{6} \gg \beta^2$$

est vérifiée et l'expression de Δu^*_s se ramène à la forme suivante :

$$\Delta u^*_s \cong -\frac{3\Delta_x}{2s-1} + \frac{3\Delta_x}{2s-1} \frac{2s-2}{2s-3} = \frac{3\Delta_x}{2s-1} \frac{1}{2s-3}. \tag{7.60}$$

Si l'on tient compte que sous les hypothèses adoptées

$$\Delta d^*_{s-1} \cong -\frac{3\Delta_x}{2s-1}, \tag{7.61}$$

la formule (7.60) peut s'écrire

$$\Delta u^*_s \cong -\frac{\Delta d^*_{s-1}}{2s-3}. \tag{7.62}$$

Si l'on pose que les conditions initiales vraies sont nulles et que $x_0 = \Delta_x$, (7.62) montre que le schéma à modèle auto-adaptatif est bien moins sensible aux conditions initiales qu'un schéma « en boucle ouverte ».

La figure 7.6 donne les diagrammes de Δu^*_{s+1} (courbe *1*) et de $-\Delta d^*_s$ (courbe *2*) en fonction du nombre d'observations pour $\beta^2 = 0{,}1$; 1 ; 10. De même que sur la figure 7.5, toutes ces courbes sont des enveloppes des fonctions de réseau. La figure 7.6 montre que dans un système à modèle auto-adaptatif l'influence des conditions initiales cesse de se faire sentir dès le quatrième ou le cinquième pas.

L'allure des courbes *1* de la figure 7.6 peut s'expliquer de la façon suivante. La régulation du modèle se fait en deux étapes. La première étape, lorsque l'effet des signaux d'entrée est encore faible et l'écart entre sorties est défini presque entièrement par la différence des conditions initiales du modèle et du système, le choix du gain du modèle doit assurer l'élimination de cette différence. C'est pourquoi pendant les premiers pas de régulation, $M\{u^*_{s+1}|\bar{\mu}\}$ peut différer considérablement de $\bar{\mu}$ et, par conséquent, $|\Delta u^*_{s+1}|$ peut être plusieurs fois supérieur à $|\Delta d^*_s|$. C'est vers le cinquième ou le sixième pas, lorsque la différence entre les conditions initiales du modèle et de l'objet est presque supprimée, que commence la régulation du modèle ayant à peu près les mêmes conditions initiales que le système, régulation qui a pour but de minimiser la différence entre $\bar{\mu}$ et le gain du modèle, car c'est cette différence qui détermine l'écart entre sorties.

Reprenons l'équation (7.39). Considérons le cas où la perturbation $\vec{h}$ représente une suite de variables aléatoires indépendantes d'espérance mathématique m_h non nulle. Cette équation peut s'écrire

$$z_s = \mu \sum_{i=1}^{s} v_i + x_0 + h_s = \mu \sum_{i=1}^{s} v_i + x_0 + m_h + \xi_s, \tag{7.63}$$

où ξ_s est une variable aléatoire centrée. En posant $x_0 + m_h = x'_0$ et en appliquant les raisonnements précédents, on peut montrer que le schéma à modèle auto-adaptatif est également peu sensible à l'erreur de définition de l'espérance mathématique de la perturbation.

Ainsi, l'avantage que présente le schéma à modèle auto-adaptatif par rapport au schéma en boucle ouverte est une sensibilité plus faible aux conditions initiales et à toute sorte d'erreurs dont les erreurs de choix des paramètres et d'affichage de gains.

§ 2. Commande optimale d'un système à retard pur

Considérons le problème de synthèse d'un organe de commande optimal pour un système à retard pur en présence des perturbations aléatoires dans la chaîne de réaction [7.9]. La commande de tels systèmes s'impose dans certaines gammes de fabrication de l'industrie chimique, de la métallurgie, etc., là où il s'agit du transport des matériaux en flot. Le phénomène du retard pur ou de parcours s'observe, par exemple, lors de l'écoulement des fluides dans des canalisations ou lors du transport des matériaux dans les fours à ciment rotatifs. Parfois le retard est imputable aux appareils de mesure. La figure 7.7 représente le schéma d'un système automatique. Le signal émis par l'objet O est canalisé avec un retard pur τ_H vers l'entrée de l'organe de commande OC par la chaîne d'action H introduisant des perturbations. L'équation de l'objet est de la forme

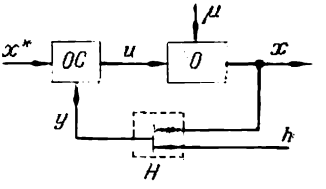


Fig. 7.7.

$$x(t+\tau_H) = \mu(t) + u(t). \tag{7.64}$$

Dans ce qui suit nous allons considérer toutes les variables à des instants discrets avec un intervalle Δt que nous allons choisir tel que $\tau = \frac{\tau_H}{\Delta t}$ soit entier. Introduisons les notations suivantes: x_s, x_s^*, u_s, μ_s pour respectivement le signal de sortie du système, la consigne, la commande et la perturbation à l'instant s; y_s pour la valeur du paramètre x_s mesuré avec une erreur h_s, c'est-à-dire

$$y_s = x_s + h_s. \tag{7.65}$$

A des instants divers, les valeurs de la perturbation sont indépendantes. Pour le temps discret, l'équation du système s'écrit

$$x_{s+\tau} = \mu_s + u_s. \tag{7.66}$$

La fonction de pertes a la forme suivante:

$$W_s = (x_s^* - x_s)^2. \tag{7.67}$$

La perturbation μ et le bruit h sont munis d'une loi normale, leurs espérances mathématiques sont nulles et leurs variances σ_μ^2 et σ_h^2.

On peut montrer, comme on l'a fait au chapitre VI, que le système automatique (7.7) est neutre. Ensuite, pour un système à retard pur, la commande u_s n'intervient que dans le risque $R_{s+\tau}$ au $(s+\tau)$-ième pas, les risques R_j, $j \neq s+\tau$, en sont indépendants.

Examinons les deux cas suivants:

1. La perturbation $\mu_s = \mu$ est une variable aléatoire.
2. La perturbation μ_s est un processus aléatoire.

Supposons que l'action perturbatrice est une variable aléatoire. En vertu des équations (7.64) et (7.65) on a alors l'égalité

$$P(y_j \mid \mu, \vec{u}_j) = P(y_j \mid \mu, u_{j-\tau}). \tag{7.68}$$

La commande optimale u_s^* est fournie en minimisant par rapport à u_s la fonction

$$\alpha_{s+\tau} = \int\limits_{\Omega(\mu)} W_s P(\mu) \prod_{j=0}^{s-1} P(y_j \mid \mu, \vec{u}_j)\, d\Omega =$$

$$= \int\limits_{-\infty}^{\infty} (x_{s+\tau}^* - \mu - u_s)^2 P(\mu) \prod_{j=0}^{s-1} P(y_j \mid \mu, u_{j-\tau})\, d\mu. \tag{7.69}$$

La densité de probabilité conditionnelle de la formule (7.69) s'écrit

$$P(y_j \mid \mu, u_{j-\tau}) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{-\frac{(y_j - \mu - u_{j-\tau})^2}{2\sigma_h^2}\right\}. \tag{7.70}$$

Après des transformations peu compliquées, on aboutit à l'expression de la commande optimale

$$u_s^* = x_{s+\tau}^* - \frac{1}{\Delta + s} \sum_{j=0}^{s-1} (y_j - u_{j-\tau}), \tag{7.71}$$

où $\Delta = \sigma_h^2/\sigma_\mu^2$. Le deuxième terme de (7.71) est l'estimation de la perturbation μ. Cette expression se distingue de la formule (6.85) par le fait qu'elle tient compte des u retardées pour un temps τ et de la valeur de la consigne dans τ pas. Si les commandes u antérieures à l'instant $s = 0$ sont inconnues et si on ne peut pas les poser égales à zéro, l'expression de u_s^* doit être modifiée de la façon suivante :

$$u_s^* = x_{s+\tau}^* - \frac{1}{\Delta + s - \tau} \sum_{j=\tau}^{s-1} (y_j - u_{j-\tau}) \quad (s > \tau). \tag{7.72}$$

Notons que les $u_0, u_1, \ldots, u_\tau$ ne sont choisies que d'après des données a priori sur la perturbation.

Etendons les résultats obtenus au cas d'un système continu (cf. § 2 du chapitre VI). Dans le cas d'un temps continu la commande optimale est donnée par la formule

$$u^*(t) = x^*(t + \tau_H) - \frac{\int\limits_0^t [y(t) - u(t - \tau_H)]\, dt}{t + a}, \tag{7.73}$$

où $a = 2\pi S_0/\sigma_\mu^2$, S_0 est la densité spectrale du bruit blanc h.

Le premier terme de la formule (7.73) peut s'obtenir à la sortie du retardateur pour un temps τ_H, le deuxième à la sortie du filtre F (fig. 7.8,a) de fonction de transition

$$\varphi(t) = \frac{t}{t+a} \tag{7.74}$$

si on applique à l'entrée du filtre la différence entre y et la commande u retardée préalablement de τ_H dans le retardateur.

En remplaçant le filtre F par un circuit apériodique à constante de temps $T = 1{,}1a$, on obtient un système voisin de l'optimal.

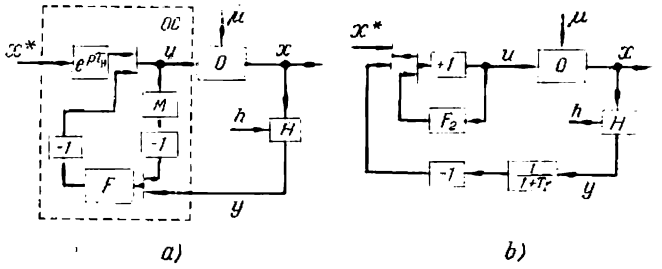


Fig. 7.8.

Comme on le voit de la figure 7.8,a, l'organe de commande optimale OC doit comporter le modèle M du système de fonction de transfert $e^{-p\tau_H}$ et le calculateur chargé à donner la moyenne statistique et les valeurs futures. Pour le système de stabilisation automatique, $x^*(t) = x^*$, et c'est pourquoi le bloc de fonction de transfert $e^{-p\tau_H}$ devient inutile. On en a tenu compte dans le schéma 7.8,b obtenu à partir du schéma 7.8,a. La fonction de transfert du filtre F_2 s'écrit

$$\Phi_2(p) = \frac{e^{-p\tau_H}}{1+Tp}. \tag{7.75}$$

On a intérêt à connaître les changements de structure dans la commande optimale en l'absence du bruit dans la chaîne de réaction. Avec $S_0 = 0$, $a = 0$, par conséquent $T = 0$. Le schéma 7.8,b se transforme en schéma 7.9,a. La fonction de transfert de l'organe de commande OC se met sous la forme

$$\Phi_3(p) = \frac{1}{1-e^{-p\tau_H}}. \tag{7.76}$$

On peut comparer un schéma construit sur la base d'un régulateur à retard de Reswick [7.10] (fig. 7.9,b) avec un schéma de régulation prévoyant l'utilisation de l'extrapolateur linéaire de Smith

[7.11] (7.9,*a*). [7.10] corcerne essentiellement le cas $K = 0$. Mais sous cette condition, les schémas 7.9,*a* et 7.9,*b* se confondent. La fonction de transfert de la partie du système comprise entre les points N_1 et N_2 (fig. 7.9,*c*) s'écrit

$$\Phi_4(p) = \frac{u(p)}{\varepsilon(p)} = \frac{1}{\dfrac{1+K_cK_vK_mG(p)}{K_cK_vG(p)} - K_m e^{-p\tau_H}}, \tag{7.77}$$

où K_v, K_m sont les gains ; $K_cG(p)$ la fonction de transfert du régulateur ; $K_cK_m(1-e^{-p\tau_H})$ celle de l'extrapolateur linéaire.

Pour $K_m = 1$ et $K_vK_c \gg 1$ on peut écrire approximativement

$$\Phi_4(p) \approx \frac{1}{1-e^{-p\tau_H}}. \tag{7.78}$$

La comparaison de (7.77) et (7.78) montre que les systèmes des figures 7.9,*a* et *c* sont identiques. L'équation du régulateur *RI* est également une approximation de l'équation (7.76) [7.9].

Ainsi le régulateur à retard de Reswick, le régulateur *RI* et le régulateur à extrapolateur linéaire de Smith doivent être considérés comme des cas particuliers d'un organe de commande optimal *OC* en l'absence de bruit dans la chaîne de réaction.

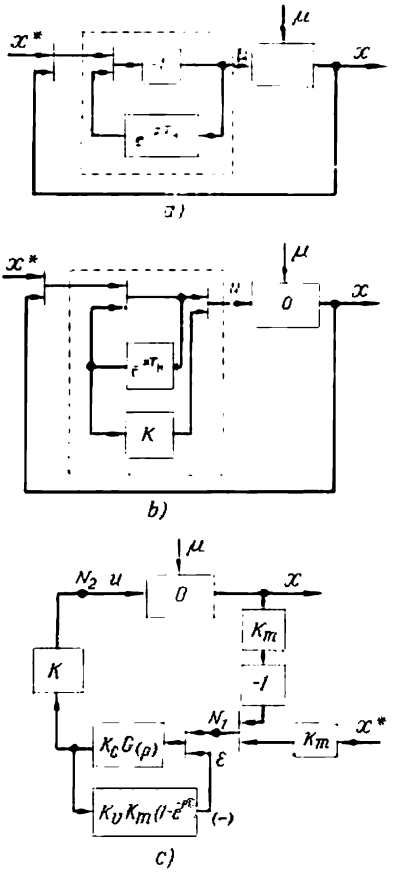


Fig. 7.9.

Recherchons la commande optimale d'un système sous la condition que la perturbation soit un processus gaussien aléatoire. De même que précédemment, commençons par le cas du temps discret, puis étendons les résultats à un système continu. Les calculs se simplifient si la fonction de corrélation normée est de la forme

$$\rho(\theta) = e^{-a_1|\theta|}, \quad 0 < a_1 < \infty. \tag{7.79}$$

Pour un processus discret le corrélogramme représente une suite de points séparés par l'intervalle Δt. Dans ce cas le processus est mar-

kovien

$$P(\vec{\mu}_s) = P(\mu_0) \prod_{j=1}^{s} P(\mu_j \mid \mu_{j-1}), \tag{7.80}$$

$$P(\mu_j \mid \mu_{j-1}) = \frac{1}{\sigma_\mu \sqrt{2\pi(1-\rho^2)}} \exp\left\{-\frac{(\mu_j - \rho\mu_{j-1})^2}{2\sigma_\mu^2(1-\rho^2)}\right\}; \tag{7.81}$$

de plus $\rho = \rho(\Delta t)$. La fonction $\alpha_{s+\tau}$ au $(s+\tau)$-ième pas est définie par l'expression

$$\alpha_{s+\tau} = \int\limits_{\Omega(\vec{\mu}_s)} (x_{s+\tau} - \mu_s - u_s)^2 P(\vec{\mu}_s) \prod_{j=0}^{s-1} P(y_j \mid \mu_{j-\tau}, u_{j-\tau})\, d\Omega. \tag{7.82}$$

La commande optimale u_s^* est définie en minimisant $\alpha_{s+\tau}$ par rapport à u_s. On a

$$u_s = x_{s+\tau}^* - \frac{\beta''_{s+\tau}}{\beta'_{s+\tau}}, \tag{7.83}$$

où

$$\beta'_{s+\tau} = \int\limits_{\Omega(\vec{\mu}_s)} P(\vec{\mu}_s) \prod_{j=0}^{s-1} P(y_j \mid \mu_{j-\tau}, u_{j-\tau})\, d\Omega, \tag{7.84}$$

$$\beta''_{s+\tau} = \int\limits_{\Omega(\vec{\mu}_s)} \mu_s P(\vec{\mu}_s) \prod_{j=0}^{s-1} P(y_j \mid \mu_{j-\tau}, u_{j-\tau})\, d\Omega, \tag{7.85}$$

d'où, compte tenu de (7.79),

$$\beta''_{s+\tau} = \rho^{\tau+1} \int\limits_{\Omega(\vec{\mu}_{s-\tau-1})} \mu_{s-\tau-1} P(\mu_0) \prod_{j=1}^{s-\tau-1} P(\mu_j \mid \mu_{j-1}) \times$$
$$\times \prod_{j=0}^{s-1} P(y_j \mid \mu_{j-\tau}, u_{j-\tau})\, d\Omega = \alpha''_{s+\tau}\rho^{\tau+1}, \tag{7.86}$$

$$\beta'_{s+\tau} = \int\limits_{\Omega(\vec{\mu}_{s-\tau-1})} P(\mu_0) \prod_{j=1}^{s-\tau-1} P(\mu_j \mid \mu_{j-1}) \times$$
$$\times \prod_{j=0}^{s-1} P(y_j \mid \mu_{j-\tau}, u_{j-\tau})\, d\Omega = d_{s+\tau}. \tag{7.87}$$

L'équation (7.83) se met alors sous la forme

$$u_s^* = x_{s+\tau}^* - \rho^{\tau+1} \frac{\alpha''_{s+\tau}}{\alpha'_{s+\tau}}. \tag{7.88}$$

En portant (7.70) et (7.81) dans (7.87), on a

$$\alpha'_{s+\tau} = A_1 \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma_h^2}\left[\mu_0^2\Delta + \sum_{j=1}^{s-\tau-1} (\mu_j - \rho\mu_{j-1})^2 \Delta' + \right.\right.$$

$$+ \sum_{j=\tau}^{s-1} (y_j - \mu_{j-\tau} - u_{j-\tau})^2 +$$

$$\left.\left. + \sum_{j=0}^{\tau-1} (y_j - \mu_{j-\tau} - u_{j-\tau})^2\right]\right\} d\mu_0\, d\mu_1 \dots d\mu_{s-\tau-1} =$$

$$= CA_1 \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma_h^2}\left[\mu_0^2\Delta - \sum_{j=1}^{s-\tau-1} (\mu_j - \rho\mu_{j-1})^2 \Delta' + \right.\right.$$

$$\left.\left. + \sum_{j=0}^{s-\tau-1} (y_{j+\tau} - \mu_j - u_j)^2\right]\right\} d\mu_0\, d\mu_1 \dots d\mu_{s-\tau-1}. \qquad (7.89)$$

Dans la formule (7.89)

$$\Delta = \frac{\sigma_h^2}{\sigma_\mu^2}, \qquad (7.90)$$

$$\Delta' = \frac{\sigma_h^2}{\sigma_\mu^2 (1-\rho^2)}, \qquad (7.91)$$

$$C = \exp\left\{-\frac{1}{2\sigma_h^2} \sum_{j=0}^{\tau-1} (y_j - \mu_{j-\tau} - u_{j-\tau})^2\right\}, \qquad (7.92)$$

$$A_1 = \frac{1}{\sigma_\mu^s \sigma_h^s (2\pi)^s (1-\rho^2)^{\frac{s-1}{2}}}. \qquad (7.93)$$

D'une manière analogue on aboutit à l'expression de $\alpha''_{s+\tau}$. Dans [7.9] est calculé le rapport $\alpha''_{s+\tau}/\alpha'_{s+\tau}$ et donnée l'expression suivante de la commande optimale :

$$u_s^* = x_{s+\tau}^* - \rho^\tau \frac{\rho}{2\sigma_h^2 p_{s-\tau-1,1}} \left[(y_{s-1} - u_{s-\tau-1}) + \frac{r}{p_{s-\tau-2,1}} \times \right.$$

$$\times \left[(y_{s-2} - u_{s-\tau-2}) + \frac{r}{p_{s-\tau-3,1}} \left[(y_{s-3} - u_{s-\tau-3}) + \dots + \right.\right.$$

$$\left.\left. + \frac{r}{p_{01}} (y_\tau - u_0)\right] \dots\right], \qquad (7.94)$$

où

$$\left.\begin{aligned} r &= \frac{\rho}{2\sigma_\mu^2 (1-\rho^2)}, \\ \frac{r}{p_{i1}} &< 1, \end{aligned}\right\} \qquad (7.95)$$

tandis que

$$\left.\begin{aligned}
&p_0 = \frac{1}{2\sigma_h^2}(1+\Delta') = p_{s-\tau-1},\\
&p_j = \frac{1}{2\sigma_h^2}(1+\Delta'+\rho^2\Delta'),\ j=1,\ 2,\ \ldots,\ s-\tau-2,\\
&p_{01} = p_0,\\
&p_{11} = p_1 - \frac{r^2}{p_{01}},\\
&\ldots\ldots\ldots\ldots\\
&p_{j1} = p_j - \frac{r^2}{p_{j-1,1}},\\
&\ldots\ldots\ldots\ldots
\end{aligned}\right\} \tag{7.96}$$

Le deuxième terme de la formule (7.94) est l'estimation de la perturbation à l'instant $s+\tau$. C'est l'espérance mathématique conditionnelle de $\mu_{s+\tau}$ pour des $\vec{y}_{s-1}$ et $\vec{u}_{s-\tau-1}$ connues. Remarquons que dans (7.94) les différences $(y_j - u_{j-\tau})$ figurent avec des poids différents: l'information plus ancienne s'altère et l'information nouvelle s'accumule. L'apparition d'un facteur devant le premier crochet de (7.94) signifie que *OC* doit prédire (dans le sens statistique) les valeurs futures de μ. En effet, pour définir l'espérance mathématique conditionnelle de $\mu_{s+\tau}$, la valeur de μ_{s-1} étant connue, il faut multiplier μ_{s-1} par le coefficient de corrélation $\rho\ (\tau\Delta t + \Delta t) = \rho^{\tau+1}$.

Passons au cas continu. Examinons d'abord un système discret à très petit pas Δt. (7.96) entraîne pour $i \to \infty$

$$p_\infty = \frac{1+\Delta'+\Delta'\rho^2}{2\sigma_h^2} - \frac{r^2}{p_\infty}, \tag{7.97}$$

puisque $p_{j1} = p_\infty \approx p_{j-1,1}$.

La résolution de l'équation quadratique (7.97) amène

$$p_\infty = \frac{1+\Delta'+\Delta'\rho^2 + \sqrt{(1+\Delta'+\rho^2\Delta') - 4\rho^2\Delta'^2}}{4\sigma_h^2} = \frac{\delta}{4\sigma_h^2}. \tag{7.98}$$

On a pris dans le numérateur de (7.98) le signe plus parce que la condition

$$p_\infty = p_{j1} \geqslant p_0 = \frac{1+\Delta'}{2\sigma_h^2} = \frac{1+\Delta'}{4\sigma_h^2} + \frac{1+\Delta'}{4\sigma_h^2} > \frac{1+\Delta'}{4\sigma_h^2} + \frac{\rho^2\Delta'}{4\sigma_h^2} \tag{7.99}$$

doit être vérifiée.

De plus

$$\frac{r}{p_\infty} = \frac{2\rho\Delta'}{\delta} < 1. \tag{7.100}$$

Etant donné que pour $\Delta t \to 0$ on peut écrire

$$1-\rho^2(\Delta t)=1-e^{-2a_1\Delta t}\approx 2a_1\Delta t, \tag{7.101}$$

on obtient l'égalité approchée

$$\frac{\delta}{\Delta}\cong 2+2(L-a_1)\,\Delta t, \tag{7.102}$$

où

$$L=\sqrt{\frac{a_1\sigma_\mu^2}{\pi S_0}+a_1^2}. \tag{7.103}$$

En utilisant (7.102), on trouve

$$\frac{r}{p_\infty}=\frac{1}{1+L\Delta t}. \tag{7.104}$$

D'une manière analogue avec de petits Δt on calcule le facteur

$$\frac{1}{2\sigma_h^2 p_{s-\tau-1,1}}=M_1\frac{\Delta t}{1+L\Delta t}, \tag{7.105}$$

où

$$M_1=\frac{1+2L\Delta t+L^2(\Delta t)^2}{\frac{\pi S_0}{a_1\sigma_\mu^2}(L+a_1)}. \tag{7.106}$$

Si la valeur de S_0 est finie et non nulle alors que $\Delta t\to 0$, cette expression devient

$$M_1\cong\frac{a_1\sigma_\mu^2}{\pi S_0(L+a_1)}. \tag{7.107}$$

Après avoir porté (7.104), (7.105) et (7.106) dans (7.94), on a

$$u_s^*=u^*(s\Delta t)=x^*(s\Delta t+\tau\Delta t)-$$

$$-\rho^{\tau+1}M_1\sum_{\nu=1}^{s}\frac{\Delta t}{(1+L\Delta t)^\nu}\,[y(s\Delta t-\nu\Delta t)-u(s\Delta t-\tau\Delta t-\nu\Delta t)]. \tag{7.108}$$

Constatons que l'observation de la condition (7.79) donne

$$\rho^\tau=\rho(\tau\Delta t)=\rho(\tau_H). \tag{7.109}$$

Passons au cas continu. En posant $\Delta t\to 0$, on a

$$u^*(t)=x^*(t+\tau_H)-\rho(\tau_H)M_1\int_0^t e^{-L\theta}[y(t-\theta)-u(t-\tau_H-\theta)]\,d\theta=$$

$$=x^*(t+\tau_H)-\rho(\tau_H)\frac{M_1}{L}L\int_0^t e^{-L\theta}[y(t-\theta)-u(t-\tau_H-\theta)]\,d\theta, \tag{7.110}$$

où $\theta = \nu\,\Delta t$. (7.110) montre que pour obtenir la commande optimale la différence entre y et la valeur de u retardée de τ_H doit être appliquée à l'entrée d'une chaîne composée d'un circuit apériodique à constante de temps $T = \frac{1}{L}$ et à gain $\xi_1 = \frac{M_1}{L}$ et d'un circuit sans inertie à gain $\xi = \rho(\tau_H) = e^{-a_1\tau_H}$. Le schéma fonctionnel de l'organe de commande optimal OC représenté sur la figure 7.10 comporte le modèle M du système de fonction de transfert $e^{-p\tau_H}$ et le calculateur effectuant les opérations suivantes sur les signaux de sortie du modèle et du système : a) mise en moyenne par le circuit apériodique et b) prédiction statistique (anticipation de la valeur de la perturbation μ dans τ_H) par le bloc ξ_2. En outre, OC doit être muni d'un bloc de prédiction des valeurs de la consigne. Pour un système de stabilisation ce bloc devient inutile puisque $x^*(t) = x^*(t + \tau_H) = x^*$.

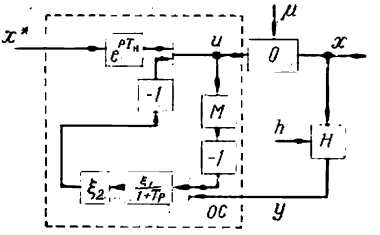


Fig. 7.10.

La comparaison des schémas (7.8,a et 7.10) montre que la structure de l'organe de commande optimal des systèmes à retard pur est la même que la perturbation μ soit une variable ou une fonction temporelle aléatoire, seuls les paramètres de OC diffèrent. En l'absence de bruits dans la chaîne de réaction, on a $S_0 = 0$, $T = 0$ et $\xi_1 = 1$, i.e. OC ne comporte plus de bloc de prise de la moyenne des valeurs mesurées du paramètre de sortie.

On a intérêt à définir l'influence qu'exerce, dans le cas d'un OC optimal, la grandeur du retard pur τ_H sur la variance σ_x^2 du paramètre de sortie $x(t)$ du système ou, ce qui revient au même, sur le risque élémentaire stable, puisque le système est linéaire et la fonction de pertes, quadratique. Supposons que la fonction d'autocorrélation de la perturbation ait la forme (7.79) et $x^*(t) = \text{const}$. En omettant les calculs intermédiaires, donnons l'expression définitive de la variance du paramètre de sortie :

$$\sigma_x^2 = \frac{\pi S_0 \xi_1^2}{T} e^{-2a_1\tau_H} + \sigma_\mu^2 \left[1 - \frac{(2-\xi_1)\,\xi_1}{1+Ta_1} e^{-2a_1\tau_H}\right], \qquad (7.111)$$

de plus

$$\xi_1 = \frac{a_1\sigma_\mu^2}{\left(a_1\sqrt{\pi S_0} + \sqrt{a_1\sigma_\mu^2 + a_1^2\pi S_0}\right)\sqrt{a_1\sigma_\mu^2 + a_1^2\pi S_0}}, \qquad (7.112)$$

$$T = \left(\frac{a_1\sigma_\mu^2}{\pi S_0} + a_1^2\right)^{-\frac{1}{2}}. \qquad (7.113)$$

Pour un système sans bruits dans la chaîne de réaction, la variance σ_x^2 peut s'écrire

$$\sigma_x^2 = \sigma_\mu^2 (1 - e^{-2a_1\tau_H}). \tag{7.114}$$

La figure 7.11 matérialise la relation entre le rapport σ_x^2/σ_μ^2 et le retard τ_H pour les cas suivants :

$S_0 = 0$; $a_1 = 0{,}5$ (courbe *1*),
$S_0 = 0{,}3\sigma_\mu^2$; $a_1 = 0{,}5$ (courbe *2*),
$S_0 = 0$; $a_1 = 0{,}05$ (courbe *3*),
$S_0 = 0{,}25\sigma_\mu^2$; $a_1 = 0{,}05$ (courbe *4*).

D'après le diagramme, plus le retard pur est petit et la pente de la courbe de la fonction d'autocorrélation de la perturbation μ douce, plus la variance est faible. Notons que pour $\tau_H \geqslant 2/a_1$

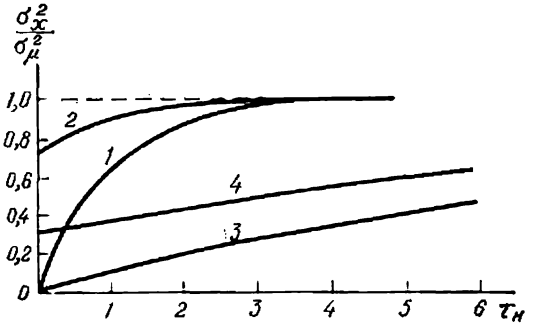


Fig. 7.11.

l'introduction d'une réaction et la construction d'un système de régulation ne donnent aucun effet. L'apparition de la perturbation dans la chaîne de réaction compromet la régulation et accroît la variance du paramètre de sortie. Un organe de commande, autre que l'organe optimal, donne naturellement lieu à une variance σ_x^2 plus grande.

§ 3. Commande avec facteurs indirects *)

Considérons le schéma du système automatique de la figure 7.12. L'objet O est composé de deux parties *1* et *2* soumises aux perturbations z_1 (μ, k) et z_2 (λ, k), fonction des paramètres aléatoires μ et λ. Désignons par f et w respectivement les variables intermédiaires non contrôlables et contrôlables. Le critère d'optimalité du processus est assujetti au signal de sortie x et à la consigne x^*.

*) L'application des méthodes de la théorie de la commande duale à certains problèmes de la commande optimale prévoyant l'utilisation des facteurs indirects est décrite dans [7.12-7.14].

Pour certaines usines de l'industrie chimique, de la métallurgie et de la cimenterie, la sortie x n'est mesurable qu'aux instants discrets séparés par un grand intervalle. La cause en est les particularités du processus de fabrication ou l'absence des capteurs et des appareils de mesure automatique continue de nombreux paramètres technologiques importants. La qualité du produit fini est évaluée

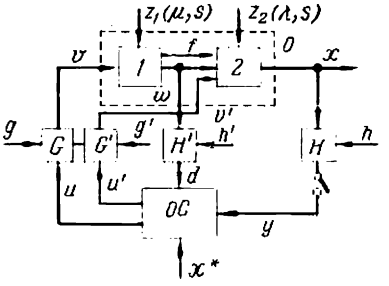


Fig. 7.12.

notamment sur la base des données fournies par les analyses immédiates effectuées en laboratoire. On contrôle automatiquement seulement certaines variables intermédiaires w (facteurs indirects) associées statistiquement au signal de sortie du système. Les paramètres w et x sont mesurés avec des erreurs aléatoires h' et h. Les commandes u et u' sont également entachées de perturbations aléatoires g et g', c'est-à-dire que les chaînes d'action G, G', H, H' entre l'organe de commande OC et le système O donnent lieu à des bruits. On considère que la combinaison des signaux et des bruits dans les chaînes et l'opérateur du système sont connus. Tous les éléments aléatoires et inconnus du système sont rapportés aux perturbations z_1 et z_2.

Comment utiliser l'information fournie par les mesures discrètes très espacées dans le temps, comment la combiner à l'information continue et aux renseignements antérieurs sur le système ? Le problème se résout par les méthodes de la théorie de la commande duale. Pour ne pas compliquer les calculs, nous allons examiner le cas unidimensionnel.

Considérons d'abord un système de commande simple (fig. 7.13). Toutes les variables sont observées à des instants discrets 0, 1, 2, . . ., k, . . ., n pris avec un intervalle Δt (n étant fixé). Affectons de l'indice k la valeur du signal au k-ième instant.

La présence d'un facteur aléatoire z_2 (λ, k) rend possible la description statistique de la relation entre x et la variable intermédiaire w. Le paramètre de sortie x n'est mesurable qu'aux instants 0, τ, 2τ, . . ., $\tau \geqslant 1$, c'est-à-dire séparés de grands intervalles

multiples de τ. Supposons que le système et les chaînes d'action H et H' sont sans mémoire. Les densités de probabilité a priori de tous les signaux aléatoires sont données. Le mode de combinaison des signaux et des bruits dans les chaînes d'action est connu, c'est-à-dire que les densités de probabilité conditionnelles $P(d_k \mid w_k)$ et $P(y_k \mid x_k)$ sont données. Les perturbations h et h' sont indépendantes l'une de l'autre et leurs valeurs à des instants voisins sont des variables indépendantes. Elles ne dépendent non plus ni de x ni de w.

Il faut définir une stratégie de l'organe de commande OC telle qu'elle minimise le risque total qui est la somme des espérances

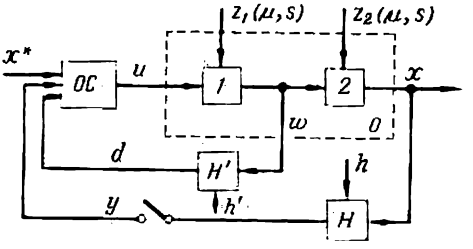


Fig. 7.13.

mathématiques d'une certaine fonction de pertes $W_k(x_k^*, k, x_k)$, le signal x^* étant connu.

D'après la méthode exposée au chapitre VI, écrivons l'expression du risque élémentaire R_k au k-ième pas. On a

$$R_k = M\{W_k \mid x_k^*\} = \int\limits_{\Omega(\mu, \lambda, \vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1})} W_k(x_k^*, k, \mu, \lambda, u_k) \times$$

$$\times P(\mu, \lambda, \vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1})\, d\Omega, \quad (7.115)$$

où

$$W_k[x_k^*, k, x_k(\mu, \lambda, u_k)] = W_k(x_k^*, k, \mu, \lambda, u_k)$$

est la fonction de pertes élémentaire au k-ième pas et

$$\vec{u}_k = (u_0, u_1, \ldots, u_k),$$

$$\vec{d}_{k-1} = (d_0, d_1, \ldots, d_{k-1}),$$

$$\vec{y}_{k-1} = (y_0, y_1, \ldots, y_{k-1})$$

les vecteurs temporels.

Utilisons également la notation

$$\vec{y}_i^* = (y_0, y_\tau, y_{2\tau}, \ldots, y_i),$$

i étant l'instant qui précède le k-ième, celui où l'on mesure le signal x. Naturellement, i est multiple de τ et $i < k$.

Pour une densité de probabilité conjointe $P(\mu, \lambda, \vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1})$ avec μ et λ indépendants on peut écrire :

$$P(\mu, \lambda, \vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1}) = P(\mu, \lambda)\, P(\vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1} \mid \mu, \lambda) = \\ = P(\mu)\, P(\lambda)\, P(\vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1} \mid \mu, \lambda). \quad (7.116)$$

D'après le théorème connu du calcul des probabilités

$$P(\vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1} \mid \mu, \lambda) = \\ = P(u_k \mid \mu, \lambda, \vec{u}_{k-1}, \vec{d}_{k-1}, \vec{y}_{k-1})\, P(\vec{u}_{k-1}, \vec{d}_{k-1}, \vec{y}_{k-1} \mid \mu, \lambda). \quad (7.117)$$

Mais $\vec{d}_{k-1}$, $\vec{u}_{k-1}$ et $\vec{y}_{k-1}$ étant fixées, la densité de probabilité de u_k ne change pas si l'on fixe encore μ et λ, puisque u_k ne dépend de μ et λ que par $\vec{d}_{k-1}$, $\vec{y}_i^*$. Par conséquent,

$$P(u_k \mid \mu, \lambda, \vec{u}_{k-1}, \vec{d}_{k-1}, \vec{y}_{k-1}) = P(u_k \mid \vec{u}_{k-1}, \vec{d}_{k-1}, \vec{y}_i^*) = \Gamma_k, \quad (7.118)$$

de plus Γ_k est la stratégie de l'organe de commande au k-ième pas.

Transformons le deuxième facteur de la formule (7.117) :

$$P(\vec{u}_{k-1}, \vec{d}_{k-1}, \vec{y}_{k-1} \mid \mu, \lambda) = P(u_0, d_0, y_0 \mid \mu, \lambda) \times \\ \times P(u_1, d_1, y_0 \mid \mu, \lambda, u_0, d_0, y_0) \ldots P(u_{\tau-1}, d_{\tau-1}, \\ y_0 \mid \mu, \lambda, \vec{u}_{\tau-2}, \vec{d}_{\tau-2}, y_0)\, P(u_\tau, d_\tau, y_\tau \mid \mu, \lambda, \vec{u}_{\tau-1}, \\ \vec{d}_{\tau-1}, y_0)\, P(u_{\tau+1}, d_{\tau+1}, y_\tau \mid \mu, \lambda, \vec{u}_\tau, \vec{d}_\tau, y_\tau) \ldots \quad (7.119)$$

L'expression (7.119) comporte deux types de facteurs

a) si $j \neq 0, \tau, 2\tau, \ldots$ et $0 < j \leqslant k-1$, on a

$$P(u_j, d_j, y_j \mid \mu, \lambda, \vec{u}_{j-1}, \vec{d}_{j-1}, \vec{y}_m^*) = \\ = P(u_j, d_j \mid \mu, \lambda, \vec{u}_{j-1}, \vec{d}_{j-1}, \vec{y}_m^*)\, \delta(y_j - y_m) = \\ = P(u_j \mid \mu, \lambda, \vec{u}_{j-1}, \vec{d}_{j-1}, \vec{y}_m^*) \times \\ \times P(d_j \mid \mu, \lambda, \vec{u}_j, \vec{d}_{j-1}, \vec{y}_m^*)\, \delta(y_j - y_m) = \\ = \Gamma_j P(d_j \mid \mu, u_j)\, \delta(y_j - y_m), \quad (7.120)$$

où m est la valeur maximale la plus proche de j tirée de la série $0, \tau, 2\tau, \ldots$; $m < j$;

b) si $j=0, \tau, 2\tau, \ldots$, on a

$$P(u_j, d_j, y_j \mid \mu, \lambda, \vec{u}_{j-1}, \vec{d}_{j-1}, \vec{y}^*_{j-\tau}) =$$
$$= P(u_j \mid \mu, \lambda, \vec{u}_{j-1}, \vec{d}_{j-1}, \vec{y}^*_{j-\tau}) \times$$
$$\times P(d_j \mid \mu, \lambda, \vec{u}_j, \vec{d}_{j-1}, \vec{y}^*_{j-\tau}) P(y_j \mid \mu, \lambda, \vec{u}_j, \vec{d}_j, \vec{y}^*_{j-\tau}) =$$
$$= \Gamma_j P(d_j \mid \mu, u_j) P(y_j \mid \mu, \lambda, u_j). \quad (7.121)$$

Les expressions (7.120) et (7.121) tiennent compte du fait que la densité de probabilité de u_j, pour des $\vec{u}_{j-1}$, $\vec{d}_{j-1}$, $\vec{y}^*_m$ fixés, ne change pas si en plus on fixe μ et λ. Ensuite, pour un système sans inertie et μ et u_j fixés, d_j ne dépend pas de λ, $\vec{u}_{j-1}$, $\vec{d}_{j-1}$, $\vec{y}^*_m$. D'une manière analogue, avec $j=0, \tau, 2\pi, \ldots$, et pour des μ, λ, u_j fixés, y_j ne dépend pas de $\vec{u}_{j-1}$, $\vec{d}_j$, $\vec{y}^*_{j-\tau}$.

Compte tenu de (7.120) et (7.121) transformons l'expression (7.119) de la façon suivante :

$$P(\vec{u}_{k-1}, \vec{d}_{k-1}, \vec{y}_{k-1} \mid \mu, \lambda) =$$
$$= \prod_{j=0}^{k-1} \Gamma_j \prod_{j=0}^{k-1} P(d_j \mid \mu, u_j) \prod_{\substack{j=1 \\ j \neq \tau, 2\tau, \ldots}}^{k-1} \delta(y_j - y_m) \times$$
$$\times \prod_{j=0, \tau, \ldots}^{i} P(y_j \mid \mu, \lambda, u_j). \quad (7.122)$$

En portant (7.118) et (7.122) dans (7.117) et le résultat obtenu dans (7.116), on obtient

$$P(\mu, \lambda, \vec{u}_k, \vec{d}_{k-1}, \vec{y}_{k-1}) = P(\mu) P(\lambda) \prod_{j=0}^{k} \Gamma_j \prod_{j=0}^{k-1} P(d_j/\mu, u_j) \times$$
$$\times \prod_{\substack{j=1 \\ j \neq \tau, 2\tau}}^{k-1} \delta(y_j - y_m) \prod_{j=0, \tau \ldots}^{i} P(y_j \mid \mu, \lambda, u_j). \quad (7.123)$$

En portant (7.123) dans la formule du risque élémentaire (7.115) et en intégrant par rapport à y_j avec $j \neq 0, \tau, 2\tau, \ldots$, on a

$$R_k = \int_{\Omega(\mu, \lambda, \vec{u}_k, \vec{d}_{k-1}, \vec{y}^*_i)} W_k(x^*_k, k, \mu, \lambda, u_k) P(\mu) P(\lambda) \times$$
$$\times \prod_{j=0}^{k} \Gamma_j \prod_{j=0}^{k-1} P(d_j \mid \mu, u_j) \prod_{j=0, \tau, \ldots}^{i} P(y_j \mid \mu, \lambda, u_j) \, d\Omega. \quad (7.124)$$

L'expression (7.124) diffère de la formule usuelle du risque élémentaire : l'expression sous le signe somme contient le produit des densités de probabilité conditionnelles de y aux instants multiples de τ, c'est-à-dire qu'il existe une chaîne supplémentaire fournissant une information sur le système. Si cette chaîne n'existe pas, y_j ne dépend pas de μ, λ, u_j et l'expression (7.124) se transforme en la formule usuelle du risque élémentaire.

Recherchons la stratégie optimale de l'organe de commande à partir de la condition de minimium du risque total en n pas. Introduisons les notations

$$\beta_k = \prod_{j=0}^{k} \Gamma_j, \tag{7.125}$$

$$\alpha_k = \int\limits_{\Omega(\mu,\,\lambda)} W_k(x_k^*, k, \mu, \lambda, u_k)\, P(\mu)\, P(\lambda) \prod_{j=0}^{k-1} P(d_j \mid \mu, u_j) \times$$
$$\times \prod_{j=0,\,\tau,\,\ldots}^{i} P(y_j \mid \mu, \lambda, u_j)\, d\Omega. \tag{7.126}$$

Le risque relatif au dernier pas

$$R_n = \int\limits_{\Omega(\vec{u}_n,\,\vec{d}_{n-1},\,\vec{y}_i^*)} \Gamma_n \beta_{n-1} \alpha_n\, d\Omega. \tag{7.127}$$

En appliquant la méthode exposée au chapitre VI on peut montrer que la stratégie de l'*OC* est pure alors que la commande optimale u_n^* se calcule en minimisant α_n par rapport à u_n. Pour l'avant-dernier pas on minimise le risque $S_{n-1} = R_n + R_{n-1}$ par rapport à u_{n-1} ; de plus

$$S_{n-1} = \int\limits_{\Omega(\vec{u}_{n-1},\,\vec{d}_{n-2},\,\vec{y}_m^*)} \Gamma_{n-1}\beta_{n-2}\alpha_{n-1}\, d\Omega + \int\limits_{\Omega(\vec{u}_n,\,\vec{d}_{n-1},\,\vec{y}_i^*)} \Gamma_n \Gamma_{n-1}\beta_{n-2}\alpha_n\, d\Omega =$$
$$= \int\limits_{\Omega(\vec{u}_{n-1},\,\vec{d}_{n-2},\,\vec{y}_m^*)} \Gamma_{n-1}\beta_{n-2}\gamma_{n-1}\, d\Omega. \tag{7.128}$$

Ici m est le nombre le plus grand ($m < n - 1$) de la série $0, \tau, 2\tau, \ldots$ qui est le plus proche de $(n - 1)$.

Dans la formule (7.128)

$$\gamma_{n-1} = \alpha_{n-1} + \int\limits_{\Omega(d_{n-1},\,y_i)} \gamma_n^*\, d\Omega \tag{7.129}$$

si y est mesuré au $(n-1)$-ième instant, c'est-à-dire si le nombre $(n-1)$ est multiple de τ, et

$$\gamma_{n-1} = \alpha_{n-1} + \int\limits_{\Omega(d_{n-1})} \gamma_n^* \, d\Omega \tag{7.130}$$

si $(n-1)$ n'est pas multiple de τ ; en outre

$$\gamma_n^* = \alpha_n^* = \alpha_n (u_n^*). \tag{7.131}$$

En poursuivant on obtient pour γ_k les expressions analogues en fonction de α_k et γ_{k+1}^*. On peut montrer que toutes les stratégies Γ_k sont des fonctions δ et que les valeurs optimales de u_k^* s'obtiennent en minimisant γ_k par rapport à u_k. La recherche de la commande optimale se ramène à des minimisations successives de γ_{k+1} par rapport à u_{k+1} et à des intégrations par rapport à d_k pour des $k \neq 0$, τ, 2τ, . . . et par rapport à d_k et y_k pour des k multiples de τ.

Ainsi les difficultés de calcul sont un peu plus grandes que dans le cas d'un système ordinaire de commande duale, mais l'ordre des calculs est toujours le même.

Considérons un exemple relativement simple. Supposons qu'un système soit décrit par les équations

$$\left.\begin{aligned} v_k &= \mu + u_k, \\ x_k &= \lambda + v_k. \end{aligned}\right\} \tag{7.132}$$

La répartition des grandeurs μ, λ, h', h est normale, leurs variances respectives sont σ_μ^2, σ_λ^2, σ_1^2, σ_h^2 et leurs espérances mathématiques nulles. La fonction de pertes s'écrit

$$W_k = (x_k^* - x_k)^2 = (x_k^* - \mu - \lambda - u_k)^2. \tag{7.133}$$

Les densités de probabilité conditionnelles de d_j et y_j sont de la forme

$$P(d_j \mid \mu, u_j) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left\{ -\frac{(d_j - \mu - u_j)^2}{2\sigma_1^2} \right\}, \tag{7.134}$$

$$P(y_j \mid \mu, \lambda, u_j) = \frac{1}{\sigma_h \sqrt{2\pi}} \exp\left\{ -\frac{(y_j - \mu - \lambda - u_j)^2}{2\sigma_h^2} \right\}. \tag{7.135}$$

Le système considéré est un système réductible et neutre. C'est pourquoi la commande optimale u_k^* se définit en minimisant la fonction α_k par rapport à u_k

$$\alpha_k = C \int\limits_{\mu=-\infty}^{\infty} \cdots \int\limits_{\lambda=-\infty}^{\infty} (x_k^* - \mu - \lambda - u_k)^2 \exp\left\{ -\frac{\mu^2}{2\sigma_\mu^2} - \frac{\lambda^2}{2\sigma_\lambda^2} - \right.$$
$$\left. - \frac{1}{2\sigma_1^2} \sum_{j=0}^{k-1} (d_j - \mu - u_j)^2 - \frac{1}{2\sigma_h^2} \sum_{j=0,\,\tau,\,\ldots}^{i} (y_j - \mu - \lambda - u_j)^2 \right\} d\mu \, d\lambda, \tag{7.136}$$

où C est une constante. En égalant la dérivée $d\alpha_k/du_k$ à zéro, on obtient

$$u_k^* = x_k^* - \frac{J_\mu}{J_0} - \frac{J_\lambda}{J_0}, \tag{7.137}$$

où

$$J_0 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left\{-\mu^2\left(\frac{1}{2\sigma_\mu^2} + \frac{k}{2\sigma_1^2} + \frac{s}{2\sigma_h^2}\right) + \right.$$

$$\left. + 2\mu\left[\frac{\sum\limits_{j=0}^{k-1}(d_j - u_j)}{2\sigma_1^2} - \frac{\sum\limits_{j=0,\,\tau,\,\ldots}^{i}(y_j - u_j)}{2\sigma_h^2}\right] - \frac{1}{2\sigma_1^2}\sum_{j=0}^{k-1}(d_j - u_j)^2\right\} \times$$

$$\times \exp\left\{-\lambda^2\left(\frac{1}{2\sigma_\lambda^2} + \frac{s}{2\sigma_h^2}\right) + 2\lambda\left[\frac{1}{2\sigma_h^2}\sum_{j=0,\,\tau,\,\ldots}^{i}(y_j - u_j) - \mu\frac{s}{2\sigma_h^2}\right] - \right.$$

$$\left. - \frac{1}{2\sigma_h^2}\sum_{j=0,\,\tau,\,\ldots}^{i}(y_j - u_j)^2\right\} d\mu\, d\lambda, \tag{7.138}$$

en outre, s est le nombre de mesures de y réalisées en k pas. Les fonctions J_μ et J_λ se distinguent de J_0 par les facteurs μ et λ dans l'expression sous le signe somme. Le calcul donne

$$u_k^* = x_k^* - A(k, s)\sum_{j=0,\,\tau,\,\ldots}^{i}(y_j - u_j) - \frac{1}{b}\,\frac{\sum\limits_{j=0}^{k-1}(d_j - u_j)}{\frac{\sigma_1^2}{\sigma_\mu^2} + k + s\frac{\sigma_1^2}{\sigma_h^2}\left(1 + \frac{\sigma_\lambda^2}{\sigma_h^2}\frac{s}{b}\right)}, \tag{7.139}$$

où

$$A(k, s) = \frac{1}{\frac{\sigma_h^2}{\sigma_\lambda^2} + s} + \frac{1}{b^2}\,\frac{1}{\frac{\sigma_h^2}{\sigma_\mu^2} + \frac{\sigma_h^2}{\sigma_1^2}k + s\left(1 + \frac{\sigma_\lambda^2}{\sigma_h^2}\frac{s}{b}\right)}, \tag{7.140}$$

$$b = 1 + \frac{\sigma_\lambda^2}{\sigma_h^2}s. \tag{7.141}$$

Considérons quelques cas particuliers.

Cas 1. Supposons que la perturbation μ est inexistante, c'est-à-dire que $\mu = 0$, $\sigma_\mu^2 = 0$.

La formule (7.139) devient alors une formule ordinaire de la commande duale d'un système sans inertie

$$u_k^* = x_k^* - \frac{\sum\limits_{j=0,\,\tau,\,\ldots}^{i}(y_j - u_j)}{\frac{\sigma_h^2}{\sigma_\lambda^2 + s}} \tag{7.142}$$

En effet, dans ce cas, la mesure de la grandeur d n'apporte aucune information nouvelle sur le système.

Cas 2. Soit $\lambda = 0$ et $\sigma_\lambda^2 = 0$. Dans ces conditions la commande optimale se trouve de la façon suivante:

$$u_k^* = x_k^* - \frac{\sum\limits_{j=0}^{k-1} (d_j - u_j)}{\frac{\sigma_1^2}{\sigma_\mu^2} + k + \frac{\sigma_1^2}{\sigma_h^2} s} - \frac{\sum\limits_{j=0,\tau,\ldots}^{i} (y_j - u_j)}{\frac{\sigma_h^2}{\sigma_\mu^2} + s + \frac{\sigma_n^2}{\sigma_1^2} k}. \tag{7.143}$$

Si les variances des perturbations dans les deux chaînes H et H' sont égales, on a

$$u_k^* = x_k^* - \frac{\sum\limits_{j=0}^{k-1} (d_j - u_j) + \sum\limits_{j=0,\tau,\ldots}^{i} (y_j - u_j)}{\frac{\sigma_h^2}{\sigma_\mu^2} + k + s}. \tag{7.144}$$

Ce cas est équivalent à celui où au lieu de $k - 1$ on a réalisé $k + s - 1$ pas, puisque pour estimer la perturbation μ on prend la moyenne de $k + s$ et non de k mesures.

Généralisons le résultat obtenu aux processus markoviens μ_k et λ_k et aux systèmes à mémoire. Pour rendre l'exposé plus simple, bornons-nous au cas unidimensionnel. Dans l'ensemble le problème reste le même, mais dans la formule du risque il faut remplacer les densités de probabilité $P(\mu)$ et $P(\lambda)$ par les densités de probabilité conjointes $P(\vec{\mu}_k)$ et $P(\vec{\lambda}_k)$. On a

$$P(\vec{\mu}_k) = P(\mu_0) \prod_{j=1}^{k} P(\mu_j \mid \mu_{j-1}), \tag{7.145}$$

$$P(\vec{\lambda}_k) = P(\lambda_0) \prod_{j=1}^{k} P(\lambda_j \mid \lambda_{j-1}), \tag{7.146}$$

où $\vec{\mu}_k$ et $\vec{\lambda}_k$ sont les vecteurs temporels. N'oublions pas non plus que pour les systèmes à mémoire les densités de probabilité conditionnelles d_k et y_k dépendent dans le cas général de tous les signaux d'entrée appliqués, c'est-à-dire de $\vec{\mu}_k$, $\vec{u}_k$ et $\vec{\mu}_k$, $\vec{\lambda}_k$, $\vec{u}_k$ respectivement. La formule du risque (7.124) devient alors

$$R_k = \int\limits_{\Omega(\vec{\mu}_k, \vec{\lambda}_k, \vec{u}_k, \vec{d}_{k-1}, \vec{y}_i^*)} W_k(x_k^*, k, \vec{\mu}_k, \vec{\lambda}_k, \vec{u}_k) P(\mu_0) P(\lambda_0) \times$$
$$\times \prod_{j=1}^{k} [P(\mu_j \mid \mu_{j-1}) P(\lambda_j \mid \lambda_{j-1})] \prod_{j=0}^{k} \Gamma_j \prod_{j=0}^{k-1} P(d_j \mid \vec{\mu}_j, \vec{u}_j) \times$$
$$\times \prod_{j=0,\tau,\ldots}^{i} P(y_j \mid \vec{\mu}_j, \vec{\lambda}_j, \vec{u}_j)\, d\Omega. \tag{7.147}$$

Dans l'ensemble, la recherche de la commande optimale reste la même, l'expression de α_k se mettant dans ce cas sous la forme

$$\alpha_k = \int\limits_{\Omega(\vec{\mu}_k, \vec{\lambda}_k)} W_k(x_k^*, k, \vec{\mu}_k, \vec{\lambda}_k, \vec{u}_k) P(\mu_0) P(\lambda_0) \times$$

$$\times \prod_{j=1}^{k} [P(\mu_j \mid \mu_{j-1}) P(\lambda_j \mid \lambda_{j-1})] \prod_{j=0}^{k-1} P(d_j \mid \vec{\mu}_j, \vec{u}_j) \times$$

$$\times \prod_{j=0, \tau, \ldots}^{i} P(y_j \mid \vec{\mu}_j, \vec{\lambda}_j, \vec{u}_j) d\Omega. \quad (7.148)$$

Les résultats qui viennent d'être obtenus pour un système unidimensionnel peuvent être étendus aux systèmes à plusieurs entrées et sorties.

Revenons au schéma 7.12 en considérant que le système et les chaînes d'action sont sans mémoire. La méthode ci-dessus conduit à la formule du risque élémentaire au k-ième pas

$$R_k = \int\limits_{\Omega(\mu, \lambda, \vec{u}_k, \vec{u}'_k, v_k, v'_k, \vec{d}_{k-1}, \vec{y}_i^*)} W_k P(\mu) P(\lambda) P(v_k \mid u_k) P(v'_k \mid u'_k) \times$$

$$\times \prod_{j=0}^{k-1} P(d_j \mid \mu, j, u_j) \prod_{j=0, \tau, .}^{i} P(y_j \mid \mu, \lambda, j, u_j, u'_j) \prod_{j=0}^{k} \Gamma_j d\Omega. \quad (7.149)$$

On peut montrer que dans ce cas-là également la stratégie optimale est pure et que les fonctions de décision Γ_k doivent être des fonctions δ, c'est-à-dire

$$\Gamma_k = P(U_k \mid \vec{u}_{k-1}, \vec{u}'_{k-1}, \vec{d}_{k-1}, \vec{y}_i^*, x^*) = \delta(U_k - U_k^*), \quad (7.150)$$

où $U_k = (u_k, u'_k)$ est le vecteur de l'espace à deux dimensions.

La recherche de la commande optimale $U_k^* = U_k^*(\vec{u}_{k-1}, \vec{u}'_{k-1}, \vec{d}_{k-1}, y_i^*, x^*)$ consiste, comme il est d'usage dans la théorie de la commande optimale, à réaliser une série de minimisations et d'intégrations de certaines fonctions, α_k se mettant sous la forme

$$\alpha_k = \alpha_k(\vec{u}_{k-1}, \vec{u}'_{k-1}, \vec{d}_{k-1}, \vec{y}_i^*, x^*) =$$

$$= \int\limits_{\Omega(\mu, \lambda, v_k, v'_k)} W_k P(\mu) P(\lambda) P(v_k \mid u_k) P(v'_k \mid u'_k) \times$$

$$\times \prod_{j=0}^{k-1} P(d_j \mid \mu, j, u_j) \prod_{j=0, \tau, \ldots}^{i} P(y_j \mid \mu, \lambda, j, u_j, u'_j) d\Omega. \quad (7.151)$$

La méthode exposée permet d'effectuer la synthèse d'un organe de commande optimal disposant d'une information incomplète sur le système et utilisant certaines variables intermédiaires liées statistiquement au signal de sortie. Pendant le fonctionnement, cette relation est précisée par la mesure périodique du signal de sortie. Parfois on dégage sans peine dans l'organe de commande *OC* la partie *A1* qui assure la commande d'après une variable intermédiaire et la partie *A2* à réponse plus lente, qui reçoit l'information sur le signal de sortie et agit en conséquence sur la partie *A1*. L'auteur

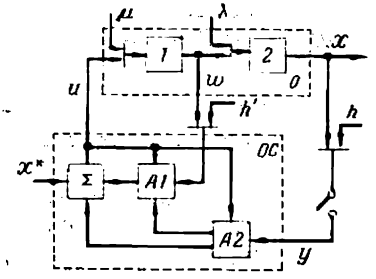


Fig. 7.14.

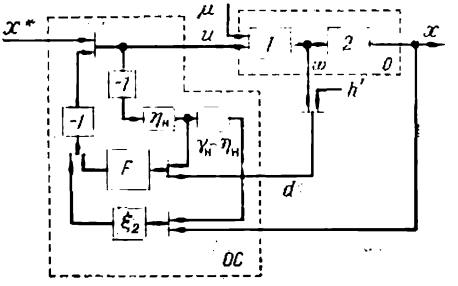


Fig. 7.15.

de [7.14] a trouvé, notamment, une structure optimale pour le système de commande à retard pur. Le schéma qu'il donne peut être ramené à celui représenté sur la figure 7.14. Σ est ici un additionneur sans inertie, *A1* comporte un filtre *F* du premier ordre aux paramètres variables qui changent par saut après l'introduction dans *A2* de la valeur suivante de y.

Cette approche permet de résoudre le problème d'un organe de commande à combinaison optimale de l'information fournie sans retard mais avec d'importantes erreurs aléatoires et de l'information obtenue avec une meilleure précision mais transmise avec un retard [7.14].

Supposons que le signal de sortie x (fig. 7.15) soit accessible à l'observation à n'importe quel instant, c'est-à-dire $\tau = 1$. La perturbation appliquée au système est un processus markovien, l'objet lui-même donnant lieu à un retard de parcours ν_H. Pour un retard important, il se peut que le système de commande n'assure pas la précision requise du paramètre de sortie [7.9], même lorsque la mesure de x est exacte (c'est-à-dire $h = 0$). On peut alors améliorer la régulation en utilisant pour la commande, outre la grandeur de sortie, la grandeur intermédiaire d associée à x (statistiquement, grâce à la présence de la perturbation h') et dont le retard de parcours η_H est sensiblement plus faible. Pour simplifier, nous allons considérer dans ce qui suit que la perturbation $z_2(\lambda)$ est inexistante.

Soient les équations du système

$$x_{k+\nu} = u_k + \mu_k, \quad v_{k+\eta} = u_k + \mu_k, \tag{7.152}$$

où la perturbation μ_k est un processus gaussien discret de fonction de corrélation normée de la forme

$$\rho(\theta) = e^{-a_1|\theta|} \quad (0 < a_1 < \infty). \tag{7.153}$$

Le processus de ce type est markovien. La consigne $x_k^* = x^* =$ $=$ const. La fonction de pertes est quadratique. La commande optimale est définie par la condition du minimum de la fonction $\alpha_{k+\nu}$ par rapport à u_k, c'est-à-dire

$$\alpha_{k+\nu} = \int\limits_{\Omega(\vec{\mu}_k)} (x^* - u_k - \mu_k)^2 P(\vec{\mu}_k) \prod_{j=0}^{k-1} [P(d_j \mid \mu_{j-\eta}, u_{j-\eta}) \times$$

$$\times P(y_j \mid \mu_{j-\nu}, u_{j-\nu})]\, d\Omega. \tag{7.154}$$

En posant

$$\frac{d\alpha_{k+\nu}}{du_k} = 0,$$

des transformations peu compliquées amènent

$$u_k^* = x^* - \rho^{\eta+1}(\Delta t)\frac{\alpha''_{k+\nu}}{\alpha'_{k+\nu}} = x^* - \rho^{\eta+1}\frac{\alpha''_{k+\nu}}{\alpha'_{k+\nu}}, \tag{7.155}$$

où

$$\alpha''_{k+\nu} = \int\limits_{\Omega(\vec{\mu}_{k-\eta-1})} \mu_{k-\eta-1} P(\mu_0) \prod_{j=1}^{k-\eta-1} P(\mu_j \mid \mu_{j-1}) \times$$

$$\times \prod_{j=0}^{k-1} [P(d_j \mid \mu_{j-\eta}, u_{j-\eta}) P(y_j \mid \mu_{j-\nu}, u_{j-\nu})]\, d\Omega. \tag{7.156}$$

La grandeur $\alpha'_{k+\nu}$ diffère de $\alpha''_{k+\nu}$ par l'absence du facteur $\mu_{k-\eta-1}$ dans l'expression sous le signe somme.

Soient μ, h' et h les perturbations normales d'espérances mathématiques nulles et de variances σ_μ^2, σ_1^2, σ_h^2. De plus

$$P(\mu_j \mid \mu_{j-1}) = \frac{1}{\sigma_\mu \sqrt{2\pi(1-\rho^2)}} \exp\left\{-\frac{(\mu_j - \mu_{j-1}\rho)^2}{2\sigma_\mu^2(1-\rho^2)}\right\}. \tag{7.157}$$

La fonction $\alpha''_{k+\nu}$ se définit de la façon suivante :

$$\alpha''_{k+\nu} = C \int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} \mu_{k-\eta-1} \exp\left\{\frac{\mu_0^2}{2\sigma_\mu^2} - \frac{1}{2\sigma_\mu^2(1-\rho^2)} \sum_{j=1}^{k-\eta-1} (\mu_j - \rho\mu_{j-1})^2 -\right.$$

$$-\frac{1}{2\sigma_1^2}\sum_{j=0}^{k-\eta-1}(d_{j+\eta}-\mu_j-u_j)^2-$$

$$-\frac{1}{2\sigma_h^2}\sum_{j=0}^{k-\nu-1}(y_{j+\nu}-\mu_j-u_j)^2\Big\}\,d\mu_0,\ \ldots,\ d\mu_{k-\eta-1}, \qquad (7.158)$$

où C est un facteur constant qui n'intervient pas dans le résultat final. Introduisons les notations

$$\zeta_j=y_{j+\nu}-u_j,\quad e_j=d_{j+\eta}-u_j, \qquad (7.159)$$

$$\Delta=\frac{\sigma_h^2}{\sigma_\mu^2},\quad \Delta_1=\frac{\sigma_h^2}{\sigma_\mu^2(1-\rho^2)},\quad \Delta_2=\frac{\sigma_h^2}{\sigma_1^2},\quad r=\frac{\rho\Delta_1}{2\sigma_h^2}. \qquad (7.160)$$

Après les calculs réalisés d'après la procédure exposée au § 7.2, on obtient

$$u_k^*=x^*-\rho^\eta\frac{\rho}{2\sigma_h^2P_{k-\eta-1,1}}\Big[\Delta_2e_{k-\eta-1}+\frac{r}{P_{k-\eta-2,1}}\Big[\Delta_2e_{k-\eta-2}+\ldots$$

$$\ldots+\frac{r}{P_{k-\nu,1}}\Big[\Delta_2e_{k-\nu}+\frac{r}{P_{k-\nu-1,1}}\Big[\zeta_{k-\nu-1}+\Delta_2e_{k-\nu-1}+\ldots$$

$$\ldots+\frac{r}{P_{01}}[\zeta_0+\Delta_2e_0]\ldots\Big]. \qquad (7.161)$$

On peut montrer que tous les facteurs r/P_{j1} de (7.161) sont inférieurs à l'unité ; de plus les P_{j1} se calculent d'après les formules de récurrence

$$\left.\begin{aligned}
P_{01}&=\frac{1}{2\sigma_h^2}(1+\Delta_1+\Delta_2),\\
P_j&=\frac{1}{2\sigma_h^2}(1+\Delta_1+\Delta_2+\rho^2\Delta_1)-\frac{r^2}{P_{j-1,1}}\\
&\qquad\text{avec } j=1,\ 2,\ \ldots,\ k-\nu-1,\\
P_{j1}&=\frac{1}{2\sigma_h^2}(\Delta_1+\Delta_2+\rho^2\Delta_1)-\frac{r^2}{P_{j-1,1}}\\
&\qquad\text{avec } j=k-\nu,\ \ldots,\ k-\eta-2,
\end{aligned}\right| \qquad (7.162)$$

$$P_{k-\eta-1}=\frac{\Delta_1+\Delta_2}{2\sigma_h^2}-\frac{r^2}{P_{k-\eta-2,1}}.$$

L'étude d'un système en régime stationnaire et l'extrapolation des résultats obtenus au cas continu ($\Delta t\to 0$) est d'un intérêt certain. Supposons que le bruit h est inexistant, alors que h' est un bruit blanc de densité spectrale

$$S'=\frac{\sigma_1^2}{2\pi}\Delta t\neq 0.$$

(7.161) permet de tirer

$$u_k^* = x^* - \rho^\eta \frac{\rho}{2\sigma_1^2 P_{k-\eta-1,1}} \left[e_{k-\eta-1} + \frac{r}{P_{k-\eta-2,1}} \times \right.$$
$$\left. \times \left[e_{k-\eta-2} + \ldots + \frac{r}{P_{k-\nu,1}} e_{k-\nu} \right] \ldots \right] - \rho^{\eta+1} C_1 P_{k-\nu-1}, \quad (7.163)$$

où C_1 est le coefficient qui dépend des caractéristiques statistiques des signaux aléatoires et de la valeur du retard.

Pour obtenir la commande optimale $u^*(t)$ dans le cas continu on recourt à la formule donnée par [7.14] :

$$u^*(t) = x^* - \rho(\eta_H) C_1 L \int_0^{\nu_H - \eta_H} [d(t - \nu_H + \eta_H + \vartheta) - u(t - \nu_H + \vartheta)] \times$$
$$\times \operatorname{sh} L\vartheta \, d\vartheta - \rho(\eta_H) C_1 [y(t) - u(t - \nu_H)], \quad (7.164)$$

où

$$L = \sqrt{\frac{a_1 \sigma_\mu^2}{\pi S'}}.$$

L'organe de commande optimal (fig. 7.15) comporte le filtre F, l'amplificateur sans inertie ξ_2 à gain $\xi_2 < 1$ et les blocs de retard de η_H et de ν_H. En remplaçant le filtre par un circuit inertiel on obtient un système de commande voisin de l'optimal.

§ 4. Systèmes d'auto-apprentissage et commande duale

La figure (7.16) représente le schéma fonctionnel d'un système auto-adaptatif [7.15, 7.16]. L'objet commandé B de sortie $\bar{x}$ (dans le cas général, $\bar{x}$ est le vecteur de coordonnées $x_1, \ldots, x_n$) forme avec l'organe de commande A_1 le système de commande principal. Dans le cas général, le système B subit la commande $\bar{u}$ et la perturbation aléatoire $\bar{z}$. La consigne $\bar{x}^*$, qui fixe les valeurs de $\bar{x}$ à atteindre à la sortie, et la grandeur $\bar{x}$ elle-même, canalisée par la chaîne de réaction, sont fournies à l'entrée de l'organe de commande A_1.

Il se peut que la commande par A_1 ne soit pas optimale, soit parce que l'information sur le système commandé B est incomplète, soit parce qu'on n'a pas muni A_1 de moyens optimaux de traitement de cette information. Dans les systèmes auto-adaptatifs, l'organe de commande comprend, en plus de la partie A_1, la partie A_2. Cette dernière peut être conçue sous la forme d'un organe autonome. Souvent les parties A_1 et A_2 sont physiquement inséparables ; elles

peuvent être, par exemple, les parties différentes d'un même programme. Toutefois, dans le schéma fonctionnel il est avantageux de les représenter comme des organes isolés. En exerçant sur A_1 l'action $\bar{w}$, la partie A_2 tend à modifier l'algorithme de A_1 pour rapporcher le système $A_1 - B$ de l'optimal. Dans certains cas on peut mesurer les perturbations $\bar{z}$ et canaliser le résultat vers les parties A_1 et A_2. Ce cas ne nous intéresse pourtant pas.

La particularité essentielle des systèmes d'auto-apprentissage, une sous-classe des systèmes auto-adaptatifs, est que le processus

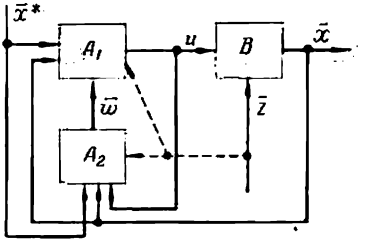


Fig. 7.16.

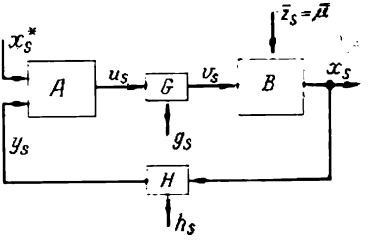


Fig. 7.17.

d'amélioration de l'algorithme de la partie A_1, de l'« éducation » de cet organe, est transitoire, alors que dans d'autres types de systèmes auto-adaptatifs c'est un processus permanent d'adaptation de la partie A_1 aux nouvelles valeurs des caractéristiques de l'objet commandé B.

La division, ne serait-ce qu'imaginaire, de l'organe de commande en parties A_1 et A_2 présente un avantage lorsque A_2 modifie assez lentement l'algorithme de A_1 à mesure de l'accumulation de l'information sur le fonctionnement du système. Par conséquent, A_1 et A_2 se distinguent foncièrement par leurs objectifs et leurs rythmes. La généralisation du schéma 7.16 permet d'imaginer une certaine hiérarchie d'organes de commande A_1, A_2, A_3, . . . dans laquelle chaque élément suivant agit sur le précédent en accumulant une information toujours plus générale sur le fonctionnement du système. C'est pourquoi dans de tels systèmes, les rythmes ralentissent avec l'ascension hiérarchique des organes.

La théorie statistique de la synthèse des systèmes optimaux qui sont dans un certain sens des systèmes d'auto-apprentissage ne s'appliquait jusqu'à présent que dans l'identification des images [7.17 à 7.19] et dans les jeux des automates [7.20]. On a établi ainsi certains algorithmes optimaux de l'apprentissage. Des algorithmes optimaux pour des types plus complexes de systèmes d'auto-apprentissage en boucle fermée peuvent s'obtenir en recourant aux méthodes de la théorie de la commande duale.

Le schéma fonctionnel 7.17 représente un système d'auto-apprentissage que nous allons étudier et dont les parties A_1 et A_2 forment encore un seul organe A. Toutes les grandeurs ne sont envisagées qu'à des instants discrets $t = 0, 1, \ldots, s, \ldots$ Une grandeur quelconque qui se manifeste à l'instant s est affectée de l'indice s (u_s, v_s, x_s, etc.). La commande u_s est amenée au système B par la chaîne sans inertie G soumise à un bruit aléatoire g_s. On connaît la loi associant la sortie v_s de la chaîne aux entrées g_s et u_s. La sortie x_s du système B est également canalisée vers l'organe de commande par la chaîne sans inertie H de bruit h_s. La sortie y_s de cette chaîne est considérée comme une fonction connue de ses entrées h_s et x_s. La consigne x_s^*, connue, est amenée directement à l'entrée de A. Le problème plus complexe de transmission de x_s^* par une chaîne avec des perturbations peut être étudié de même qu'au § 3 du chapitre V.

Pour simplifier, supposons que le système B soit sans inertie. La généralisation aux systèmes avec inertie peut être effectuée par les méthodes décrites au § 4 du chapitre VI. Le système subit la perturbation $\bar{z}_s = \bar{\mu}$, où $\bar{\mu}$ est généralement un vecteur. La relation entre la sortie x_s du système B et les entrées v_s et $\bar{\mu}$ est

$$x_s = F(v_s, \bar{\mu}). \tag{7.165}$$

La perturbation $\bar{\mu}$ est un vecteur aléatoire de densité de probabilité a priori $P_0(\bar{\mu} \mid \bar{a})$ dépendant d'un certain vecteur des paramètres $\bar{a}$ de coordonnées $a_1, \ldots, a_m$. Supposons que les paramètres de la loi de répartition P_0 sont inconnus avant le lancement du système. On ne connaît que leur densité de probabilité a priori conjointe $q_0(a_1, \ldots, a_m) = q_0(\bar{a})$. Supposons ensuite que les valeurs de $\bar{\mu}$ diffèrent d'un cycle à l'autre des processus de la même série, mais que les paramètres a_i aient certaines valeurs constantes bien qu'inconnues a priori. La connaissance de l'estimation approchée de ces paramètres contribue à l'amélioration du processus de commande. Plus le nombre de cycles de la série est grand, plus les estimations des paramètres sont précises, ce qui permet également de mieux estimer la perturbation $\bar{\mu}$.

Dans divers cycles ($j = 1, \ldots, N$) de la même série, la grandeur $\bar{\mu}$ peut prendre des valeurs différentes $\bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_k, \ldots, \bar{\mu}_N$ que nous considérons comme des échantillons indépendants d'une même population de même valeur de vecteur $\bar{a}$. Dans les séries différentes, le vecteur $\bar{a}$ peut être différent, il constitue alors une variable aléatoire de densité de probabilité $q_1(\bar{a})$. C'est pourquoi la densité a priori $P_0(\bar{\mu})$ s'obtient par la mise en moyenne de $P_0(\bar{\mu} \mid \bar{a})$ par rapport au vecteur $\bar{a}$ pour différentes séries d'expériences :

$$P_0(\bar{\mu}) = \int_{\Omega(\bar{a})} P_0(\bar{\mu} \mid \bar{a})\, q_0(\bar{a})\, d\Omega. \tag{7.166}$$

Ici $\Omega(\bar{a})$ est le domaine de variation du vecteur $\bar{a}$ et $d\Omega$ son élément infiniment petit.

Nous considérons que les perturbations $\bar{\mu}$, g_s et h_s sont indépendantes; g_s de même que h_s sont supposées être des suites des variables indépendantes de densité de probabilité $P(g_s)$ et $P(h_s)$ respectivement.

Supposons que chaque série donne lieu à N cycles de commande, chaque cycle étant composé de $(n+1)$ pas qui sont des instants discrets $t = 0, 1, \ldots, n$ (chaque cycle a sa numération de t). N et n sont fixés. Introduisons la fonction de pertes élémentaire pour le t-ième pas du j-ième cycle:

$$W_{sj} = W_s(s, x_{sj}, x_s^*). \tag{7.167}$$

Les fonctions W_s et x_s^* aux instants $t = s$ sont les mêmes pour tous les cycles. On désigne par x_{sj} la valeur de x au s-ième pas du j-ième cycle.

Considérons que le système optimal est tel qu'il minimise le risque moyen R

$$R = M\{W\} = M\{\sum_{j=1}^{N} \sum_{s=0}^{n} W_s(s, x_{sj}, x_s^*)\} = \min. \tag{7.168}$$

Dans ces conditions, les commandes u_{sj} doivent appartenir à un certain domaine admissible $\Omega(u)$.

Le problème consiste à rechercher la stratégie optimale de l'organe de commande A qui peut en général être aléatoire et qui constitue une densité conditionnelle Γ_{sk} de la grandeur u_{sk} au s-ième pas du k-ième cycle. Γ_{sk} peut être fonction des variables appliquées dans le passé à l'entrée de A. x_s^* étant considérée connue, on peut ne pas l'expliciter. Par conséquent, Γ_{sk} dépend généralement de y_{ij} et u_{ij} $(i = 0, \ldots, n;\ j = 1, \ldots, k-1)$, ainsi que de y_{ik} et u_{ik} avec $i = 0, \ldots, s-1$. Introduisons les vecteurs variables

$$\vec{u}_{sj} = (u_{0j}, \ldots, u_{sj}),\quad \vec{y}_{sj} = (y_{0j}, \ldots, y_{sj}),\quad \vec{x}_{sj} = (x_{0j}, \ldots, x_{sj}), \tag{7.169}$$

ainsi que les matrices de ces vecteurs, $\vec{\vec{u}}_{sj}$ par exemple:

$$\vec{\vec{u}}_{sj} = \begin{vmatrix} \vec{u}_{sj} \\ \vec{u}_{s,\,j-1} \\ \vdots \\ \vec{u}_{s1} \end{vmatrix} = \begin{vmatrix} u_{0j}, & \ldots, & u_{js} \\ u_{0,\,j-1}, & \ldots, & u_{s,\,j-1} \\ \vdots & & \\ u_{01}, & \ldots, & u_{s1} \end{vmatrix}. \tag{7.170}$$

Dans le cas général, Γ_{sk} dépend alors des matrices $\vec{\vec{y}}_{n,\,k-1}$, $\vec{\vec{u}}_{n,\,k-1}$ qui contiennent les variables correspondantes des cycles antiérieurs

de la série considérée, ainsi que des vecteurs $\vec{y}_{s-1,\,k}$ et $\vec{u}_{s-1,\,k}$ qui contiennent les valeurs de ces mêmes variables aux pas antérieurs du cycle considéré :

$$\Gamma_{sk} = \Gamma_{sk}\,(u_{sk} \mid \vec{y}_{s-1,\,k},\ \vec{u}_{s-1,\,k},\ \vec{\bar{y}}_{n,\,k-1},\ \vec{\bar{u}}_{n,\,k-1}). \qquad (7.171)$$

On demande de rechercher les fonctions optimales Γ_{sk} ($s = 0, \ldots, n$; $k = 1, \ldots, N$) qui minimisent le risque moyen R. Γ_{sk} étant la densité de probabilité, elle doit être une fonction non négative vérifiant la condition

$$\int\limits_{\Omega(u_{sk})} \Gamma_{sk}\,(u_{sk})\,d\Omega = 1. \qquad (7.172)$$

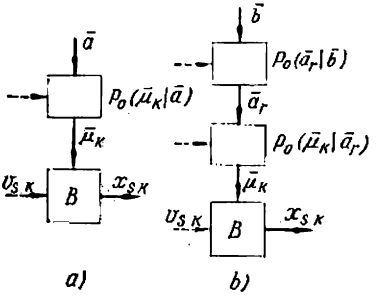


Fig. 7.18.

Cette position du problème est une certaine généralisation de certains problèmes de la théorie de la commande duale dans lesquels la densité de probabilité a priori de $\bar{\mu}$ était considérée comme connue. Dans la théorie des systèmes auto-adaptatifs on a souvent affaire à des situations à forte indétermination [2.38]. On peut étudier ces situations par des voies différentes, en appliquant, par exemple, les critères du minimax. Le mode indiqué ci-dessus d'« aggravation » de l'indétermination en renonçant à fixer le vecteur $\bar{a}$ et à donner la densité $q_0\,(\bar{a})$ constitue peut-être la généralisation la plus naturelle de la méthode bayesienne fort utilisée dans la statistique mathématique. L'interprétation immédiate en est donnée par un schéma équivalent du système B (fig. 7.18). $\bar{\mu}_k$ est ici une sortie du bloc à mécanisme aléatoire (ou à entrée supplémentaire sous la forme d'une variable représentée en pointillé) munie de la loi $P_0\,(\bar{\mu}_k \mid \bar{a})$ (fig. 7.18, *a*). La grandeur $\bar{a}$ de densité de probabilité a priori $q_0\,(\bar{a})$ donnée est appliquée à l'entrée du bloc. L'« aggravation » ultérieure de l'indétermination peut s'obtenir en adjoignant encore un bloc (fig. 7.18, *b*). La grandeur $\bar{a}_r$ peut constituer la sortie du bloc à mécanisme aléatoire muni de la loi $P_0\,(\bar{a}_r \mid \bar{b})$. La variable $\bar{b}$ de densité de probabilité a priori donnée est appliquée à l'entrée de ce bloc. Chaque nouveau bloc supplémentaire augmente l'indétermination.

Déduisons d'abord l'expression du risque élémentaire conditionnel r_{sk} au s-ième pas du k-ième cycle en entendant par r_{sk} l'espérance mathématique de la fonction de pertes élémentaire W_{sk}, la « préhistoire » des entrées de l'organe de commande A notée $I_{s-1,\,k}$

étant donnée :

$$I_{s-1,\,k} = (\vec{y}_{n,\,k-1},\ \vec{u}_{n,\,k-1},\ \vec{y}_{s-1,\,k},\ \vec{u}_{s-1,\,k}). \tag{7.173}$$

On obtient l'expression

$$r_{sk} = M\{W_{sk} \mid I_{s-1,\,k}\} = \int\limits_{\Omega(x_{sk})} W_s(s,\ x_{sk})\, P(x_{sk} \mid I_{s-1,\,k})\, d\Omega =$$

$$= \int\limits_{\Omega(x_{sk}\bar{\mu}_k,\ u_{sk})} W_s(s,\ x_{sk})\, P(x_{sk} \mid \bar{\mu}_k,\ u_{sk})\, P_{sk}(\bar{\mu}_k) \times$$

$$\times\, \Gamma_{sk}(u_{sk} \mid I_{s-1,\,k})\, d\Omega. \tag{7.174}$$

Désignons par $P_{sk}(\bar{\mu}_k)$ la densité de probabilité a posteriori

$$P_{sk}(\bar{\mu}_k) = P(\bar{\mu}_k \mid I_{s-1,\,k}). \tag{7.175}$$

Au lieu de $\bar{\mu}_k$ introduisons d'abord la matrice $\bar{\bar{\mu}}_k$ des vecteurs $\bar{\mu}_j$ $(j = 1,\ \ldots,\ k)$. Ecrivons la densité de probabilité conjointe de $\bar{\bar{\mu}}_k$ et de $\bar{a}$:

$$P(\bar{\bar{\mu}}_k,\ \bar{a}) = q_0(\bar{a})\, P(\bar{\bar{\mu}}_k \mid \bar{a}) = q_0(\bar{a}) \prod_{j=1}^{k} P_0(\bar{\mu}_j \mid \bar{a}) \tag{7.176}$$

(cette dernière transformation étant vérifiée en vertu de l'indépendance des $\bar{\mu}_j$ pour différents j), la densité de probabilité a priori s'écrit alors

$$P_0(\bar{\bar{\mu}}_k) = \int\limits_{\Omega(\bar{a})} P(\bar{\bar{\mu}}_k,\ \bar{a})\, d\Omega = \int\limits_{\Omega(\bar{a})} q_0(\bar{a}) \prod_{j=1}^{k} P_0(\bar{\mu}_j \mid \bar{a})\, d\Omega. \tag{7.177}$$

La densité de probabilité a posteriori de $\bar{\bar{\mu}}_k$ s'obtient à partir de la formule de Bayes

$$P(\bar{\bar{\mu}}_k \mid I_{s-1,\,k}) = \frac{P(I_{s-1,\,k} \mid \bar{\bar{\mu}}_k)\, P_0(\bar{\bar{\mu}}_k)}{P(I_{s-1,\,k})}, \tag{7.178}$$

où $P(I_{s-1,\,k})$ est la densité de probabilité a priori de $I_{s-1,\,k}$. En portant (7.177) dans (7.178) on obtient l'expression de la densité de probabilité a posteriori (7.175) du vecteur $\bar{\mu}_k$:

$$P(\bar{\mu}_k \mid I_{s-1,\,k}) = \int\limits_{\Omega(\mu_{k-1})} P(\bar{\bar{\mu}}_k \mid I_{s-1,\,k})\, d\Omega =$$

$$= \int\limits_{\Omega(\bar{\mu}_2,\ \ldots,\ \bar{\mu}_{k-1},\ \bar{a})} q_0(\bar{a}) \left[\prod_{j=1}^{k} P_0(\bar{\mu}_j \mid \bar{a})\right] \frac{P(I_{s-1,\,k} \mid \bar{\bar{\mu}}_k)}{P(I_{s-1,\,k})}\, d\Omega. \tag{7.179}$$

Développons l'expression $P(I_{s-1,k}\,|\,\bar{\bar{\mu}}_k)$ sous le signe somme:

$$P(I_{s-1,k}\,|\,\bar{\bar{\mu}}_k) = P(\vec{\vec{y}}_{n,k-1}, \vec{\vec{y}}_{n,k-1}, \vec{y}_{s-1,k}, \vec{u}_{s-1,k}\,|\,\bar{\mu}_1, \ldots, \bar{\mu}_k) =$$
$$= P(\vec{y}_{n1}, \vec{u}_{n1}\,|\,\bar{\mu}_1)\, P(\vec{y}_{n2}, \vec{u}_{n2}\,|\,\vec{y}_{n1}, \vec{u}_{n1}, \bar{\mu}_1, \bar{\mu}_2) \times \ldots$$
$$\ldots \times P(\vec{y}_{nj}, \vec{u}_{nj}\,|\,\vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}, \bar{\mu}_1, \ldots, \bar{\mu}_j) \times \ldots$$
$$\ldots \times P(\vec{y}_{n,k-1}, \vec{u}_{n,k-1}\,|\,\vec{\vec{y}}_{n,k-2}, \vec{\vec{u}}_{n,k-2}, \bar{\mu}_1, \ldots, \bar{\mu}_{k-1}) \times$$
$$\times P(\vec{y}_{s-1,k}, \vec{u}_{s-1,k}\,|\,\vec{\vec{y}}_{n,k-1}, \vec{\vec{u}}_{n,k-1}, \bar{\mu}_1, \ldots, \bar{\mu}_k). \quad (7.180)$$

Le facteur type (sauf le dernier) s'écrit alors

$$P(y_{0j}, y_{1j}, \ldots, y_{nj};\ u_{0j}, u_{1j}, \ldots, u_{nj}\,|\,\vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1},$$
$$\bar{\mu}_1, \ldots, \bar{\mu}_j) = P(y_{0j}, u_{0j}\,|\,\vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}, \bar{\mu}_1, \ldots, \bar{\mu}_j) \times$$
$$\times P(y_{1j}, u_{1j}\,|\,y_{0j}, u_{0j}, \vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}, \bar{\mu}_1, \ldots, \bar{\mu}_j) \cdot P(y_{2j}, u_{2j}\,|\,\vec{y}_{1j},$$
$$\vec{u}_{1j}, \vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}, \bar{\mu}_1, \ldots, \bar{\mu}_j) \ldots P(y_{nj}, u_{nj}\,|\,\vec{y}_{n-1,j},$$
$$\vec{u}_{n-1,j}, \vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}, \bar{\mu}_1, \ldots, \bar{\mu}_j). \quad (7.181)$$

A son tour le facteur type de la dernière expression peut se mettre sous la forme

$$P(y_{ij}, u_{ij}\,|\,\vec{y}_{i-1,j}, \vec{u}_{i-1,j}, \vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}, \bar{\mu}_1, \ldots, \bar{\mu}_j) =$$
$$= P(u_{ij}\,|\,\vec{y}_{i-1,j}, \vec{u}_{i-1,j}, \vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}) \cdot P(y_{ij}\,|\,u_{ij}, \bar{\mu}_j) =$$
$$= \Gamma_{ij} \cdot P(y_{ij}\,|\,u_{ij}, \bar{\mu}_j). \quad (7.182)$$

L' expression est vraie parce que la densité de probabilité de y_{ij} est complètement définie en fixant u_{ij} et $\bar{\mu}_j$ et ne change pas avec une information supplémentaire sur $\vec{y}_{i-1,j}$, $\vec{u}_{i-1,j}$, $\vec{\vec{y}}_{n,j-1}$, $\vec{\vec{u}}_{n,j-1}$, $\bar{\mu}_1, \ldots, \bar{\mu}_{j-1}$.

Ainsi

$$P(\vec{y}_{nj}, \vec{u}_{nj}\,|\,\vec{\vec{y}}_{n,j-1}, \vec{\vec{u}}_{n,j-1}, \bar{\mu}_1, \ldots, \bar{\mu}_j) =$$
$$= \{\prod_{i=0}^{n} \Gamma_{ij}\} \{\prod_{i=0}^{n} P(y_{ij}\,|\,u_{ij}, \bar{\mu}_j)\}. \quad (7.183)$$

Un raisonnement analogue conduit à l'expression

$$P(\vec{y}_{s-1,k}, \vec{u}_{s-1,k}\,|\,\vec{\vec{y}}_{n,k-1}, \vec{\vec{u}}_{n,k-1}, \bar{\mu}_1, \ldots, \bar{\mu}_k) =$$
$$= \{\prod_{i=0}^{s-1} \Gamma_{ik}\} \{\prod_{i=0}^{s-1} P(y_{ik}\,|\,u_{ik}, \bar{\mu}_k)\}. \quad (7.184)$$

(7.182)-(7.184) entraînent

$$P(I_{s-1,k}\,|\,\bar{\bar{\mu}}_k) = P(\vec{\bar{y}}_{n,k-1}, \vec{\bar{u}}_{n,k-1}, \vec{y}_{s-1,k}, \vec{u}_{s-1,k}\,|\,\bar{\mu}_1, \ldots, \bar{\mu}_k) =$$
$$= \prod_{j=1}^{k-1} \{(\prod_{i=0}^{n} \Gamma_{ij}) \prod_{i=0}^{n} P(y_{ij}\,|\,u_{ij}, \bar{\mu}_j)\} \times$$
$$\times (\prod_{i=0}^{s-1} \Gamma_{ik}) \{\prod_{i=0}^{s-1} P(y_{ik}\,|\,u_{ik}, \bar{\mu}_k)\}. \quad (7.185)$$

En portant (7.185) dans (7.179) et (7.179) dans (7.174) on tombe sur la formule de r_{sk}. En prenant la moyenne de r_{sk} sur l'ensemble des expériences nous obtenons l'expression de R_{sk}

$$R_{sk} = \int_{\Omega(I_{s-1,k})} r_{sk} P(I_{s-1,k})\, d\Omega =$$
$$= \int_{\Omega(\bar{\bar{\mu}}_k, \bar{a}, x_{sk}, \vec{\bar{y}}_{n,k-1}, \vec{\bar{u}}_{n,k-1}, \vec{y}_{s-1,k}, \vec{u}_{sk})} W_s(s, x_{sk}) \times$$
$$\times P(x_{sk}\,|\,\bar{\mu}_k, u_{sk})\, q_0(\bar{a}) \,[\prod_{j=1}^{k} P_0(\bar{\mu}_j\,|\,\bar{a})] \prod_{j=1}^{k-1} \{(\prod_{i=0}^{n} \Gamma_{ij}) \times$$
$$\times \prod_{i=0}^{n} P(y_{ij}\,|\,u_{ij}, \bar{\mu}_j)\} (\prod_{i=0}^{s} \Gamma_{ik}) \cdot \prod_{i=0}^{s-1} P(y_{ik}\,|\,u_{ik}, \bar{\mu}_k)\, d\Omega. \quad (7.186)$$

Désignons par ρ_{k-1} l'expression

$$\rho_{k-1}(\bar{a}, \vec{\bar{y}}_{n,k-1}), \vec{\bar{u}}_{n,k-1}) =$$
$$= \int_{\Omega(\bar{\bar{\mu}}_{k-1})} q_0(\bar{a}) \prod_{j=1}^{k-1} \Big[P_0(\bar{\mu}_j\,|\,\bar{a}) \Big] \cdot \prod_{j=1}^{k-1} \Big\{ \prod_{i=0}^{n} P(y_{ij}\,|\,u_{ij}, \bar{\mu}_j) \Big\}\, d\Omega. \quad (7.187)$$

Il vient

$$R_{sk} = \int_{\Omega(\bar{a}, \bar{\mu}_k, x_{sk}, \vec{y}_{s-1,k}, \vec{u}_{sk}, \vec{\bar{y}}_{n,k-1}, \vec{\bar{u}}_{n,k-1})} W_s(s, x_{sk}) \times$$
$$\times P(x_{sk}\,|\,\bar{\mu}_k, u_{sk})\, P_0(\bar{\mu}_k\,|\,\bar{a})\, \rho_{k-1} \Big(\prod_{j=1,\, i=0}^{j=k-1,\, i=n} \Gamma_{ij} \Big) \times$$
$$\times \Big(\prod_{i=0}^{s} \Gamma_{ik} \Big) \prod_{i=0}^{s-1} P(y_{ik}\,|\,u_{ik}, \bar{\mu}_k)\, d\Omega. \quad (7.188)$$

L'expression de ρ_{k-1} contient l'information fournie par les observations réalisées pendant les cycles précédents. Notons

$$\varphi_j(\bar{a}, \vec{y}_{n,j}, \vec{u}_{n,j}) = \int\limits_{\Omega(\bar{\mu}_j)} P_0(\bar{\mu}_j \mid \bar{a}) \cdot \prod_{i=0}^{n} P(y_{ij} \mid u_{ij}, \bar{\mu}_j)\, d\Omega. \tag{7.189}$$

On en tire

$$\rho_{k-1}(\bar{a}, \vec{\vec{y}}_{n,k-1}, \vec{\vec{u}}_{n,k-1}) = q_0(\bar{a}) \prod_{j=1}^{k-1} \varphi_j(\bar{a}, \vec{y}_{nj}, \vec{u}_{nj}). \tag{7.190}$$

On peut montrer que la grandeur ρ_{k-1} est proportionnelle à la densité de probabilité a posteriori $P(\bar{a} \mid I_{0,k-1})$ et contient une information plus riche sur le vecteur inconnu des paramètres $\bar{a}$.

Dans le cas particulier où la densité a priori $\bar{\mu}_j$ ne dépend pas de $\bar{a}$, il est aisé de montrer que tout le processus se décompose en cycles indépendants. Toutefois, si la densité $\bar{\mu}_j$ dépend de $\bar{a}$, alors, comme on le montre dans ce qui suit, le comportement optimal au j-ième cycle est défini par l'information obtenue aux cycles antérieurs tout en tenant compte des processus éventuels des cycles ultérieurs.

Procédons maintenant au choix d'une stratégie optimale.

Considérons d'abord le risque élémentaire R_{nN} au dernier pas ($t = n$) du dernier cycle ($j = N$). Posons

$$\beta_{N-1} = \prod_{j=1,\, i=0}^{j=N-1,\, i=n} \Gamma_{ij}, \quad \beta_{n-1,N} = \prod_{i=0}^{n-1} \Gamma_{iN}. \tag{7.191}$$

Il vient

$$R_{nN} = \int\limits_{\Omega(\bar{a}, \bar{\mu}_N, x_{nN}, \vec{y}_{n-1,N}, \vec{u}_{nN}, \vec{\vec{y}}_{n,N-1}, \vec{\vec{u}}_{n,N-1})} W_n(n, x_{nN}) \times$$
$$\times P(x_{nN} \mid \bar{\mu}_N, u_{nN}) P_0(\bar{\mu}_N \mid \bar{a}) \rho_{N-1} \beta_{N-1} \beta_{n-1,N} \prod_{i=0}^{n-1} P(y_{iN} \mid u_{iN}, \bar{\mu}_N)\, d\Omega. \tag{7.192}$$

Soit

$$\alpha_{nN}(\vec{y}_{n-1,N}, \vec{u}_{nN}, \vec{\vec{y}}_{n,N-1}, \vec{\vec{u}}_{n,N-1}) = \int\limits_{\Omega(\bar{a}, \bar{\mu}_N, x_{nN})} W_n(n, x_{nN}) \times$$
$$\times P(x_{nN} \mid u_{nN}, \bar{\mu}_N) \cdot P_0(\bar{\mu}_N \mid \bar{a}) \cdot \rho_{N-1}(\bar{a}, \vec{\vec{y}}_{n,N-1}, \vec{\vec{u}}_{n,N-1}) \times$$
$$\times \prod_{i=0}^{n-1} P(y_{iN} \mid u_{iN}, \bar{\mu}_N)\, d\Omega. \tag{7.193}$$

On en tire

$$R_{nN} = \int\limits_{\Omega(\vec{y}_{n-1,N},\, \vec{u}_{nN},\, \vec{\vec{y}}_{n,N-1},\, \vec{\vec{u}}_{n,N-1})} \beta_{N-1}\cdot\beta_{n-1,N}\cdot\alpha_{nN}\Gamma_{nN}\, d\Omega =$$

$$= \int\limits_{\Omega(\vec{y}_{n-1,N},\, \vec{u}_{n-1,N},\, \vec{y}_{n,N-1},\, \vec{\vec{u}}_{n,N-1})} \beta_{N-1}\cdot\beta_{n-1,N}\cdot\varkappa_{nN}\, d\Omega, \qquad (7.194)$$

où

$$\varkappa_{nN} = \int\limits_{\Omega(u_{nN})} \alpha_{nN}\Gamma_{nN}\, d\Omega. \qquad (7.195)$$

Pour minimiser R_{nN} il faut choisir u_{nN} telle que, quels que soient $\vec{y}_{n-1,N}$, $\vec{u}_{n-1,N}$, $\vec{\vec{y}}_{n,N-1}$, $\vec{\vec{u}}_{n,N-1}$, la fonction $\varkappa_{nN}$ ait une valeur minimale.

On en tire la stratégie optimale Γ_{nN}. La valeur optimale u^*_{nN} s'obtient en minimisant α_{nN} par rapport à u_{nN}. Dans ces conditions, la commande optimale u^*_{nN} devient fonction de l'information fournie antérieurement à l'organe de commande A

$$u^*_{nN} = u^*_{nN}(\vec{y}_{n-1,N},\ \vec{u}_{n-1,N},\ \vec{\vec{y}}_{n,N-1},\ \vec{\vec{u}}_{n,N-1}), \qquad (7.196)$$

de plus

$$\alpha^*_{nN} = \min_{u_{nN}} (\alpha_{nN}) = \alpha_{nN}(u^*_{nN}). \qquad (7.197)$$

En rétrogradant pas à pas dans le dernier cycle, puis en passant à l'avant-dernier cycle etc., on aboutit à la stratégie optimale obtenue de la façon suivante : introduisons la fonction

$$\alpha_{sk} = \alpha_{sk}(\vec{\vec{y}}_{n,k-1},\ \vec{\vec{u}}_{n,k-1},\ \vec{y}_{s-1,k},\ \vec{u}_{s-1,k}) =$$

$$= \int\limits_{\Omega(\bar{a},\, x_{sk},\, \bar{\mu}_k)} W_s(s, x_{sk})\, P(x_{sk} \mid \bar{\mu}_k, u_{sk})\, P_0(\bar{\mu}_k \mid \bar{a}) \times$$

$$\times \rho_{k-1}\cdot \prod_{i=0}^{s-1} P(y_{ik} \mid u_{ik}, \bar{\mu}_k)\, d\Omega. \qquad (7.198)$$

Introduisons également la fonction γ_{sk}, de plus $\gamma_{nN} = \alpha_{nN}$, et la valeur minimale de γ_{sk} se note γ^*_{sk}. L'expression de γ_{sk} s'obtient à partir des relations de récurrence :

$$\gamma_{sk} = \alpha_{sk} + \int\limits_{\Omega(y_{sk})} \gamma^*_{s+1,k}\, d\Omega \qquad (0 \leqslant s < n,\ 1 \leqslant k \leqslant N), \qquad (7.199)$$

$$\gamma_{nk} = \alpha_{nk} + \int\limits_{\Omega(y_{nk})} \gamma^*_{0,k+1}\, d\Omega \qquad (k < N).$$

La stratégie optimale élémentaire Γ^*_{sk} est une stratégie pure, c'est-à-dire qu'elle a la forme d'une fonction δ

$$\Gamma^*_{sk} = \delta(u_{sk} - u^*_{sk}), \tag{7.200}$$

où u^*_{sk} résulte de la minimisation de γ_{sk} par rapport à u_{sk}

$$\gamma^*_{sk} = \min_{\Omega(u_{sk})} \gamma_{sk} = (\gamma_{sk})_{u_{sk}=u^*_{sk}}. \tag{7.201}$$

Evidemment, pour $s > 0$, $1 < k \leqslant N$ on a la relation

$$u^*_{sk} = u^*_{sk}(\vec{y}_{s-1,k}, \vec{u}_{s-1,k}, \vec{\vec{y}}_{n,k-1}, \vec{\vec{u}}_{n,k-1}), \tag{7.202}$$

et pour $k = 1$, $0 < s \leqslant n$

$$u^*_{s1} = u^*_{s1}(\vec{y}_{s-1,1}, \vec{u}_{s-1,1}). \tag{7.203}$$

Enfin pour $k = 1$, $s = 0$, la grandeur $u^*_{0,1}$ se calcule à partir des considérations a priori. Il faut alors poser

$$\prod_{i,j} P(y_{ij} \mid u_{ij}, \bar{\mu}_1) = 1.$$

La stratégie définie par les conditions (7.198)-(7.203) est typique pour la commande duale. En effet, la vitesse d'apprentissage caractérisée par la densité de probabilité a posteriori $P_{sk}(\bar{\mu}_k)$ dépend des stratégies élémentaires Γ_{ij} utilisées par l'organe de commande. C'est pourquoi, dans le cas général, la grandeur u^*_{sk} est définie non seulement par la condition de minimum de R_{sk} mais encore par l'influence exercée par la commande u^*_{sk} sur la vitesse d'apprentissage et, par conséquent, sur les valeurs futures des risques élémentaires.

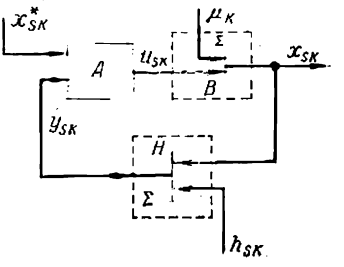


Fig. 7.19.

Examinons un exemple illustrant la théorie qui vient d'être décrite. Pour rendre l'exposé plus simple, étudions le système neutre représenté sur la figure 7.19.

L'équation du système B s'écrit

$$x_{sk} = \mu_k + u_{sk}. \tag{7.204}$$

L'équation de la chaîne d'action H est la suivante

$$y_{sk} = x_{sk} + h_{sk}. \tag{7.205}$$

Supposons que la consigne x^*_{sk} soit connue d'avance. La fonction de pertes élémentaire s'écrit alors

$$W_{sk} = (x^*_{sk} - x_{sk})^2, \tag{7.206}$$

h_{sk}, μ_k et a étant des grandeurs gaussiennes de densité de probabilité

$$\left.\begin{aligned} P(h_{sk}) &= \frac{1}{\sigma_h\sqrt{2\pi}}\exp\left\{-\frac{h_{sk}^2}{2\sigma_n^2}\right\}, \\ P_0(\mu_k \mid a) &= \frac{1}{\sigma_\mu\sqrt{2\pi}}\exp\left\{-\frac{(\mu_k-a)^2}{2\sigma_\mu^2}\right\}, \\ q_0(a) &= \frac{1}{\sigma_a\sqrt{2\pi}}\exp\left\{-\frac{(a-a_0)^2}{2\sigma_a^2}\right\}. \end{aligned}\right\} \tag{7.207}$$

L'effet d'apprentissage est grand si $\sigma_a \gg \sigma_\mu$ et il en résulte que la densité a priori $P_0(\mu_k)$ (cf. (7.166)) est bien plus « étalée » sur de différentes valeurs possibles que la densité a priori dans le cas d'un grand nombre de cycles lorsque la valeur de a est précisée.

Etant donné que

$$y_{ij} = x_{ij} + h_{ij} = u_{ij} + \mu_j + h_{ij}, \tag{7.208}$$

il vient

$$P(y_{ij} \mid u_{ij}, \mu_j) = \frac{1}{\sigma_h\sqrt{2\pi}}\exp\left\{-\frac{(y_{ij}-u_{ij}-\mu_j)^2}{2\sigma_h^2}\right\}. \tag{7.209}$$

Le calcul d'après les formules (7.189) et (7.190) conduit à l'expression

$$\rho_{k-1} = B_{k-1}\exp\{-b_{k-1}a^2 + c_{k-1}a + d_{k-1}\}, \tag{7.210}$$

où B_{k-1} est une certaine grandeur dépendant de k, et

$$\left.\begin{aligned} b_{k-1} &= \frac{1}{2\sigma_a^2} + \frac{k-1}{\vartheta\sigma_\mu^4}, \\ c_{k-1} &= \frac{a_0}{\sigma_a^2} + \frac{2}{\vartheta\sigma_\mu^2\sigma_h^2}\sum_{j=1}^{k-1}(\Sigma_j), \\ d_{k-1} &= -\frac{a_0^2}{2\sigma_a^2} + \frac{1}{\vartheta\sigma_h^4}\sum_{j=1}^{k-1}(\Sigma_j^2) - \frac{1}{2\sigma_h^2}\sum_{j=1}^{k-1}\theta_j. \end{aligned}\right\} \tag{7.211}$$

ϑ, Σ_j et θ_j s'expriment ici par les égalités

$$\vartheta = \frac{1}{2\sigma_\mu^2} + \frac{n+1}{2\sigma_h^2},$$

$$\Sigma_j = \sum_{i=0}^{n}(y_{ij}-u_{ij}), \quad \theta_j = \sum_{i=0}^{n}(y_{ij}-u_{ij})^2. \tag{7.212}$$

Les coefficients b_k se calculent pour chaque valeur de k; ils ne dépendent pas des observations. En même temps les c_k et d_k dépendent des observations réalisées par l'organe de commande A.

Les c_k et d_k peuvent se calculer à partir de c_{k-1} et d_{k-1} d'après les formules de récurrence résultant de (7.211)

$$\left.\begin{aligned} c_k &= c_{k-1} + \frac{2}{\vartheta\sigma_\mu^2\sigma_h^2}\Sigma_k, \\ d_k &= d_{k-1} + \frac{1}{\vartheta\sigma_h^4}\Sigma_k^2 - \frac{1}{2\sigma_h^2}\theta_k. \end{aligned}\right\} \tag{7.213}$$

Pour $k = 1$, c'est-à-dire dans le premier cycle, il n'existe pas encore de données fournies par les cycles antérieurs. Donc

$$\rho_0 = q_0(a) = \frac{1}{\sigma_a\sqrt{2\pi}} \exp\left\{-\frac{a^2}{2\sigma_a^2} + \frac{aa_0}{\sigma_a^2} - \frac{a_0^2}{2\sigma_a^2}\right\}. \tag{7.214}$$

On en déduit que

$$b_0 = \frac{1}{2\sigma_a^2}, \quad c_0 = \frac{a_0}{\sigma_a^2}, \quad d_0 = -\frac{a_0^2}{2\sigma_a^2}. \tag{7.215}$$

Ces expressions avec (7.213) peuvent servir pour calculer c_k et d_k. Les calculs d'après la formule (7.198) amènent l'expression

$$\alpha_{sk} = \frac{E_{sk}}{2p_{sk}}\sqrt{\frac{\pi}{p_{sk}}}\left[1 + \frac{(q_{sk} - 2p_{sk}v_{sk})^2}{2p_{sk}}\right] \exp\left\{\frac{q_{sk}^2}{4p_{sk}}\right\}, \tag{7.216}$$

où E_{sk} dépend de k et des observations précédentes (nous ne donnons pas cette fonction car nous n'en aurons pas besoin dans la suite) et les coefficients p_{sk}, v_{sk} et q_{sk} sont définis par les égalités

$$\left.\begin{aligned} p_{sk} &= \frac{1}{2\sigma_\mu^2} + \frac{s}{2\sigma_h^2} - \frac{1}{4\sigma_\mu^4\left(b_{k-1} + \frac{1}{2\sigma_\mu^2}\right)}, \\ q_{sk} &= \frac{\Sigma_{s-1,k}}{\sigma_h^2} + \frac{c_{k-1}}{2\sigma_\mu^2\left(b_{k-1} + \frac{1}{2\sigma_\mu^2}\right)}, \\ \Sigma_{s-1,k} &= \sum_{i=0}^{s-1}(y_{ik} - u_{ik}), \quad v_{sk} = x_{sk}^* - u_{sk}. \end{aligned}\right\} \tag{7.217}$$

Les expressions (7.216) et (7.217) montrent que seul le terme entre crochets de la formule (7.216) pour a_{sk} dépend de u_{sk} dissimulé dans le coefficient v_{sk}. C'est pourquoi, par exemple, la condition de minimum de a_{nN} par rapport à u_{nN} correspond à l'égalité

$$q_{nN} - 2p_{nN}v_{nN} = 0. \tag{7.218}$$

En y portant la valeur de v_{nN} on en tire

$$u_{nN}^* = x_{nN}^* - \frac{q_{nN}}{2p_{nN}}. \tag{7.219}$$

Le dernier terme est l'estimation optimale de μ_N.

L'examen de l'expression $a^*_{nN} = \gamma^*_{nN}$ montre que $y_{n-1,\,N}$ n'y fait partie que dans la différence $y_{n-1,\,N} - u_{n-1,\,N} = z$. C'est pourquoi en remplaçant dans l'intégrale $\int\limits_{-\infty}^{\infty} \gamma^*_{nN}\, dy_{n-1,\,N}$ la variable d'intégration par z on peut voir que cette intégrale ne dépend pas de $u_{n-1,\,N}$. Par conséquent, la valeur optimale $u^*_{n-1,\,N}$ s'obtient en minimisant le seul terme $\alpha_{n-1,\,N}$ par rapport à $u_{n-1,\,N}$. Un raisonnement analogue permet d'établir que, généralement, la valeur optimale u^*_{sk} s'obtient par minimisation de l'expression de a_{sk} par rapport à u_{sk}. Il en résulte donc que

$$u^*_{sk} = x^*_{sk} - \frac{q_{sk}}{2p_{sk}} = x^*_{sk} - \Bigg[\frac{a_0}{\sigma_a^2\sigma_\mu^2\nu_k L_{sk}} + \\ + \frac{2}{\sigma_h^2\sigma_\mu^2\vartheta\nu_k L_{sk}} \sum_{j=1}^{k-1}\sum_{i=0}^{n} (x_{ij} - u_{ij}) + \frac{1}{\sigma_h^2 L_{sk}} \sum_{i=0}^{s-1} (x_{ik} - u_{ik})\Bigg], \quad (7.220)$$

où

$$\nu_k = \frac{1}{\sigma_a^2} + \frac{1}{\sigma_\mu^2} + 2\,\frac{(n+1)(k-1)}{\vartheta\sigma_\mu^2\sigma_h^2}, \quad (7.221)$$

$$L_{sk} = \frac{1}{\sigma_\mu^2} + \frac{s}{\sigma_h^2} - \frac{1}{\sigma_\mu^4\nu_k}. \quad (7.222)$$

Le deuxième terme entre crochets tient compte des observations réalisées pendant les cycles précédents.

L'algorithme (7.220) peut être réalisé dans le schéma fonctionnel 7.20. Les valeurs des différences $(y_{sk} - u_{sk})$ sont canalisées jusqu'au bloc M_1 de la partie A_1 de l'organe de commande, bloc qui calcule $\Sigma_{s-1,\,k}$. Une fois le cycle achevé, l'information accumulée par le bloc en est effacée. La partie A_2 de l'organe de commande transmet au bloc L_1 les valeurs b_{k-1} et c_{k-1}. Le bloc L_1 est régi par la formule (7.220).

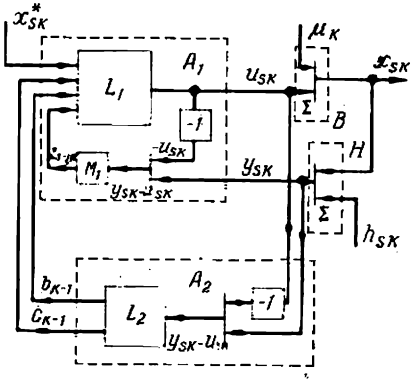


Fig. 7.20.

La partie A_2 qui emmagasine l'expérience des cycles antérieurs comporte un bloc de calcul de la différence $(y_{sk} - u_{sk})$, un bloc doublant M_1 et, enfin, un bloc de calcul de b_{k-1} et c_{k-1} d'après les formules de récurrence (7.213) et (7.215).

La division de l'organe A en parties A_1 et A_2 est dans une certaine mesure conventionnelle. On peut représenter, par exemple, le bloc sans inertie L_1 sous la forme d'une partie de l'organe de comman-

de, le bloc M_1 sous la forme d'une autre partie mémorisant l'information d'un caractère particulier pendant le cycle courant, le bloc L_2 sous la forme d'une troisième partie accumulant l'information d'un caractère plus général. Pourtant, dans ce cas également l'organe de commande dans son ensemble est une hiérarchie des parties qui accumulent une information de plus en plus générale.

On peut montrer qu'avec la croissance du numéro du cycle, le risque moyen en un cycle, c'est-à-dire la grandeur

$$S_k = \sum_{s=0}^{n} R_{sk}, \tag{7.223}$$

diminue.

En portant (7.220) dans (7.188) on obtient par intégration la valeur minimale du risque élémentaire R_{sk} :

$$R_{sk} = \sigma_h^2 + \frac{s}{\sigma_h^2 L_{sk}^2} + \frac{(n+1)(k-1)(\sigma_h^2+\sigma_\mu^2)}{(\sigma_h^2\sigma_\mu^2\vartheta\nu_k L_{sk})^2} + \\ + \sigma_\mu^2\left(\frac{s}{\sigma_h^2 L_{sk}} - 1\right)^2 + \frac{1}{(L_{sk}\nu_k\sigma_\mu^2)^2\,\sigma_a^2}. \tag{7.224}$$

Si $\sigma_a \gg \sigma_\mu$, c'est-à-dire si la variance a priori du paramètre a est grande et l'indétermination principale est due à une mauvaise

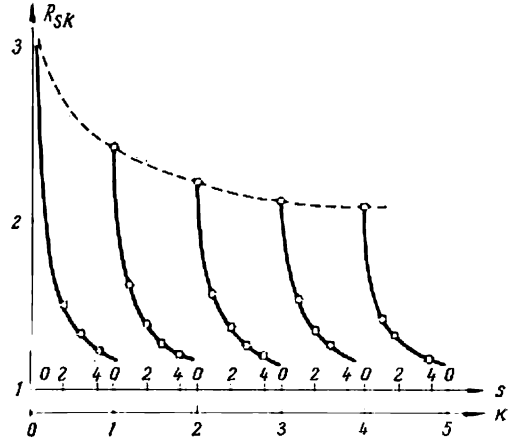


Fig. 7.21.

connaissance de ce paramètre, l'accumulation de l'information est la plus rapide. Cette inégalité permet de simplifier (7.224)

$$R_{sk} = \sigma_h^2 + \frac{s}{\sigma_h^2 L_{sk}^2} + \frac{(n+1)(k-1)(\sigma_h^2+\sigma_\mu^2)}{(\sigma_h^2\sigma_\mu^2\vartheta\nu_k L_{sk})^2} + \sigma_\mu^2\left(\frac{s}{\sigma_h^2 L_{sk}} - 1\right)^2. \tag{7.225}$$

La formule (7.225) montre que lorsque s croît, pour k fixé, R_{sk} décroît, et lorsque k croît, s étant fixé, R_{sk} décroît de façon monotone. La figure 7.21 représente la variation de R_{sk} en fonction des valeurs $\sigma_\mu = \sigma_h = 1$. Ce diagramme montre que déjà vers le cinquième cycle le système se stabilise, c'est-à-dire que le paramètre a est assez bien défini.

La figure 7.22 représente la variation du risque total S_k durant un cycle. La vitesse de décroissance de S_k d'un cycle à l'autre caracérise la vitesse du processus d'apprentissage qui consiste, dans

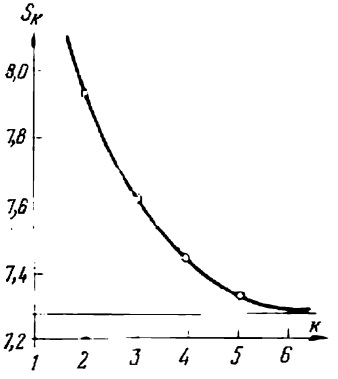


Fig. 7.22.

notre exemple, dans l'accumulation de l'expérience par la partie A_2 de l'organe de commande dans le but d'améliorer l'algorithme de la partie A_1 en modifiant le coefficient c_{k-1}. Le risque élémentaire R_{sk} pour $k = \text{const}$ diminue avec l'accroissement de k par suite l'accumulation de l'information pendant le cycle. On pourrait appeler cet effet « apprentissage en germe », car il est lié à la modification dirigée de l'algorithme du bloc L_1. Plus généralement, dans tout système optimal, l'accumulation de l'information contribue à l'apprentissage du système plus ou moins efficace. La différence entre divers types de systèmes est définie dans une grande mesure par le volume d'information accumulée et, par suite, par l'existence d'une hiérarchie plus ou moins développée des parties de l'organe de commande.

CONCLUSION

Quelles sont les tendances et les perspectives les plus proches de la théorie des systèmes optimaux ?

Il est difficile de donner à cette question une réponse complète du fait que l'essor de cette théorie peut définir dès demain des tournants imprévus et des situations nouvelles. On ne peut formuler que certaines considérations qui se présentent comme très probables.

Tout d'abord il faut fournir un grand effort pour jeter un pont entre la théorie et la pratique. Les applications pratiques de la théorie des systèmes optimaux sont encore insuffisantes bien que leur nombre ne cesse de croître. L'étude des types concrets des systèmes optimaux à utiliser dans l'industrie chimique, métallurgique et autres, la mise au point des stratégies optimales de commande des systèmes de transport et énergétiques complexes poseront plusieurs nouveaux problèmes dont la résolution favorisant l'activité pratique impulsera le développement de la théorie. La conception des systèmes voisins des optimaux et la détermination des algorithmes quasi optimaux sont un problème théorique d'actualité très important (cf., par exemple, [3.67]). Pour lui suffire, il faut mettre à profit l'arsenal richissime des méthodes générales de l'analyse moderne, de l'analyse fonctionnelle entre autres, et les moyens puissants que fournissent les techniques de calcul. Les questions relatives à la synthèse pratique des systèmes font également partie de ce corps de problèmes ; cette synthèse porte non pas sur les algorithmes ou les schémas fonctionnels abstraits, mais sur des schémas réels des organes de commande optimaux et de leurs blocs.

Le développement de la théorie des systèmes optimaux conduira à son extension au cas où l'on dispose de l'information a priori non seulement sur le système, mais encore sur l'organe de commande. La sélection de l'organe de commande sera ainsi soumise à certaines contraintes. Ce sont des problèmes qui jusqu'à présent n'ont été que peu étudiés et n'ont pas été examinés dans notre ouvrage. Or, l'importance d'une telle approche ne soulève aucun doute. Tout organe de commande réel se constitue à partir d'éléments réels aux

possibilités restreintes. Ces éléments subissent l'action des perturbations ; leur bande de fréquences ou, plus généralement, leur bande passante dans le sens informatique est bornée.

La fiabilité de ces éléments n'est pas suffisante ; ils peuvent donner lieu à des défaillances alors que l'ensemble du système doit fonctionner sans aléas. Enfin, des considérations économiques imposent le choix de la version susceptible d'assurer un avantage maximal, par exemple sur le plan des investissements ou des frais d'exploitation par unité produite. La théorie des systèmes optimaux débouche dans ce domaine sur la théorie de la fiabilité et les problèmes qu'elle pose sont très proches des problèmes économiques et autres qui appartiennent, en fait, à la théorie des systèmes complexes.

La troisième orientation de la théorie des systèmes optimaux est celle du concept d'indétermination. L'indétermination, l'absence d'une information a priori exhaustive sur le système et, d'autant plus, sur l'organe de commande, sont caractéristiques de la théorie de la commande. Dans notre ouvrage nous avons insisté à plus d'une reprise sur ce fait. Les problèmes que nous avons étudiés forment essentiellement deux groupes dont l'un intéresse les systèmes à information complète sur l'objet commandé et l'autre les systèmes à information incomplète mais disposant des renseignements a priori probabilistes (méthode bayésienne).

La réalité donne pourtant lieu à des cas très fréquents lorsque les caractéristiques probabilistes a priori sont absentes du tout ou ne sont connues que partiellement, ou encore, changent d'une façon imprévisible. Dans ces cas, l'information sur l'objet commandé au sein du système considéré est encore moins complète que dans les systèmes passés en revue dans notre ouvrage. L'extension de la théorie des systèmes optimaux aux cas qui viennent d'être indiqués est très nécessaire de nos jours du point de vue pratique. Cette extension fait l'objet d'une partie de [2.38]. Certes, la généralisation d'une théorie peut se faire par des voies différentes. Il semble que dans le domaine considéré on fera toujours plus appel à une méthode inspirée par la théorie des jeux [5.10-5.14, 5.32-5.35]. Cette méthode subira probablement des modifications au fur et à mesure que se développera la théorie des jeux qui interviendra dans des domaines toujours nouveaux, tels les jeux à plusieurs participants, pas forcément antagonistes, par exemple, pouvant coopérer provisoirement et s'entr'aider en vue d'un bénéfice commun. Le développement de la théorie des jeux des automates présente également, à notre avis, une tendance prometteuse [6.18]. Pourtant, dans le « jeu » contre la nature la méthode inspirée par la théorie des jeux n'est pas toujours assez justifiée (cf. chapitre V). C'est pourquoi on verra se développer d'autres méthodes, par exemple la méthode réunissant le principe du minimax à l'accumulation de l'information au cours du travail du système (cf. notamment, [5.36]). Cette méthode met à

profit tous les renseignements sur le système accumulés jusqu'à l'instant donné. Supposons que ces renseignements soient insuffisants pour définir d'une façon univoque les caractéristiques probabilistes et qu'il leur corresponde un ensemble de caractéristiques possibles. Pourtant, cet ensemble est borné par l'information recueillie au cours du processus et son volume est plus petit que celui de départ. Plus la quantité de l'information fournie à l'organe de commande est grande, plus le volume de l'ensemble considéré est petit. En choisissant dans cet ensemble les caractéristiques « les plus mauvaises » et les plus dangereuses sous l'optique du critère d'optimalité adopté, c'est-à-dire en appliquant le principe du minimax, on peut établir pour le cas considéré une stratégie optimale du minimax. A mesure que l'information sur le système s'accumule, cette stratégie se rapproche de plus en plus de la stratégie optimale associée à une information complète ou maximale sur le système.

Une autre approche, qui apparemment diffère peu de la précédente, consiste à appliquer le concept de probabilité induite. L'origine de cette notion remonte à Bernoulli et Laplace, fondateurs du calcul des probabilités. A notre époque, ce concept est approfondi par R. Carnap [5.37]. Certains problèmes utilisant la notion de probabilité induite sont étudiés par Bellman [2.38] et [5.38].

Voici un exposé très simple du principe de la probabilité induite. Supposons que nous ne savons rien sur un certain événement aléatoire. Nous n'avons alors aucune raison de considérer sa probabilité supérieure ou inférieure à $\frac{1}{2}$. On la prendra donc égale à $\frac{1}{2}$. D'autre part, si l'événement s'est produit dans m de n épreuves pour un n grand, la fréquence m/n peut être prise pour la mesure de la probabilité. La formule qui donne le résultat égal à $\frac{1}{2}$ pour $n = m = 0$ et le résultat égal à m/n dans le cas de m et n grands s'écrit :

$$p_i = \frac{1+m}{2+n}.$$

C'est précisément la formule la plus simple de la probabilité induite. Elle donne, à partir de l'expérience antérieure, la mesure de la probabilité future d'un événement aléatoire, le nombre d'épreuves effectuées dans le passé pouvant être quelconque, grand ou petit. Les questions relatives à la justification des formules de ce type dépassent, certes, le cadre du calcul des probabilités [5.37].

En général, l'établissement des méthodes d'exploration des situations indéterminées est d'un intérêt capital pour comprendre le fonctionnement des systèmes organisés complexes, ainsi que pour développer la théorie des systèmes auto-adaptatifs et par là même toute la cybernétique.

Une tendance importante de la théorie des systèmes optimaux est la position toujours plus universelle des problèmes et la mise au

point des méthodes de résolution de plus en plus générales. La tendance d'envisager d'un point de vue unique les problèmes les plus divers subsistera dans l'avenir et aura pour conséquence une liaison toujours plus intime de la théorie des systèmes optimaux avec d'autres orientations de la cybernétique technique.

Une telle jonction est très importante. En effet, elle garantit la formation d'une base théorique et scientifique unique pour l'ensemble de la cybernétique. Il semble qu'une telle base et les lois générales définissant les possibilités des systèmes cybernétiques reposerons sur les principes de l'informatique. En effet, les éléments fondamentaux de la cybernétique sont les concepts associés à l'information et à son traitement.

La théorie de l information brillante et profonde de Shannon avec ses ramifications actuelles est l'un des premiers pas dans la formation des principes scientifiques de la cybernétique. Le concept de la quantité d'information introduit par Shannon suffit lors d'une étude isolée des problèmes de transmission de l'information par un canal. Mais la théorie du traitement de l'information dans les systèmes en boucle fermée impose l'introduction dans l'informatique de concepts nouveaux, au moins de celui de sens et de valeur de l'information. La discussion de ces concepts dépasse le cadre du présent ouvrage. Bornons-nous à illustrer la nécessité de cette introduction sur l'exemple de l'activité des systèmes biologiques. Un homme est soumis à un grand nombre de stimuli de toutes sortes et, s'il réagissait à chacun d'eux, son activité se traduirait par des mouvements désordonnés les plus variés. Or en fait, un homme agit de façon consciente et sélectionne l'information reçue d'après son sens, c'est-à-dire d'après la relation qui existe entre l'information et les objets réels, d'après la valeur de celle-ci, c'est-à-dire d'après son utilité pour les actions envisagées par l'homme.

Les processus optimaux de la sélection de l'information, la perception, la formation des conceptions générales et, enfin, l'établissement des projets, la préparation des actions, en d'autres termes, la simulation de la pensée dans un domaine défini, autant d'éléments qui doivent trouver leur place dans la théorie générale des systèmes optimaux. Les processus d'auto-apprentissage, d'auto-organisation et d'autres types d'adaptation pourront alors être étudiés dans une optique unique pour être rangés d'après leur efficacité.

La logique de l'automatisation conduit non seulement à l'extension illimitée du domaine d'action des systèmes automatiques mais encore à la complication qualitative des problèmes qui se posent.

En effet, l'automatisation des travaux de routine, des travaux nécessitant une faible et une moyenne qualification permet à un nombre toujours plus grand d'hommes de se consacrer à une activité plus qualifiée, à caractère souvent auxiliaire. Il s'agit avant tout du contrôle, du réglage, de la planification, de la recherche, etc. L'au-

tomatisation ultérieure vise à aider l'homme dans ses activités hautement intellectuelles. L'extension des automates technologiques assurant le processus de production proprement dit s'accompagnera de celle des automates de contrôle, de réglage, de planification ; on verra apparaître des automates qui faciliteront la recherche et remplaceront l'homme à certaines de ses étapes.

Bien entendu, le remplacement de plus en plus poussé de l'homme par les automates ne risque pas de l'évincer de la production. C'est l'homme qui commandera toujours un système automatique et plus ce dernier se complique, plus important devient le rôle de l'homme.

Les automates assument les fonctions de plus en plus complexes, y compris les fonctions relatives à la conception des systèmes de commande. Une question se pose à savoir quel sera, aux nouvelles étapes de l'automatisation, le rôle de la théorie, y compris de la théorie des systèmes optimaux ? Les nouvelles tendances de l'automatisation ne risquent-elles pas de remplacer l'homme par les automates dans le domaine de l'élaboration de la stratégie optimale ?

Supposons que le couplage des blocs élémentaires d'un organe de commande soit d'abord aléatoire; en cours d'apprentissage et d'adaptation le schéma fonctionnel et ses paramètres changent automatiquement jusqu'à ce qu'on arrive à l'optimum. Quel est le rôle de la théorie des systèmes optimaux ? Ne verrons-nous pas son déclin après avoir assisté à son essor ? Ne deviendra-t-elle pas inutile dans l'avenir ?

Cette question se rapporte dans la même mesure à toute théorie. Selon nous il ne peut y avoir qu'une réponse; jamais une théorie ne peut être remplacée par l'expérimentation tâtonnante, même effectuée par des automates et non par des hommes. La science et la technique de l'avenir auront encore plus besoin de la pensée, des hypothèses, des constructions, donc de la théorie. C'est pourquoi la thèse du dépérissement de la théorie est fausse. Cependant la théorie doit changer de caractère. Si un automate effectue des calculs plus vite, d'une façon plus simple et avec une précision plus grande qu'un homme, ces calculs doivent être réalisés par l'automate. Même si la tâche du théoricien se limite à la présentation du problème sous la forme d'un programme à traiter sur une machine, jusqu'à ce stade le champ d'activité de la théorie est très vaste.

Par ailleurs, l'interaction homme-machine pose de nouveaux problèmes théoriques. Comment devraient être les blocs élémentaires d'un automate d'auto-apprentissage pour qu'il assure la résolution de problèmes suffisamment nombreux ? Quelles sont les conditions de convergence du processus d'apprentissage ? Quelle est la relation entre les caractéristiques du processus d'apprentissage et l'état initial de l'automate ? Comment assurer l'approximation optimale en cours d'apprentissage ? Enfin, que faire pour rendre optimal le

processus d'apprentissage, par exemple, comment rendre minimale sa durée ou obtenir le résultat le plus précis possible ?

Ainsi la nouvelle étape de l'automatisation fait naître des problèmes théoriques plus généraux que les précédents. Leur résolution impose de nouvelles méthodes, de nouvelles approches. Plus encore, l'emploi des machines étend les possibilités de la théorie et met à l'ordre du jour les problèmes qui hier encore étaient inimaginables. En définitive, l'interaction des machines et de la théorie contribue à leur perfectionnement et à leur développement.

Les problèmes qui se posent à la théorie se compliquent en même temps que l'activité pratique de l'homme exige des connaissances d'un niveau plus élevé. Cet essor permanent est la condition sine qua non de la maîtrise de la technique actuelle et de la création de la technique des lendemains.

BIBLIOGRAPHIE

Chapitre premier

1.1. WIENER N. — Cybernetics. New York, John Wiley, 1948.

1.2. FELDBAUM A. — Optimal Systems. Dans le recueil de PESCHON J. Disciplines and technics of systems control. Blaisdell Publishing Compagny, New York, 1965.

1.3. KHARKÉVITCH A. — Sur une réponse en fréquence spéciale pour évaluer la distorsion d'une courbe (en russe). « Journal de Physique théorique », 1936, t. 6, fasc. 9, p. 1551-1960.

1.4. KRASSOVSKI A. — Critères intégraux d'un processus de régulation (en russe). Comptes rendus du séminaire sur la théorie de l'asservissement. Moscou, Machguiz, 1949.

1.5. NEWTON G. C., GOULD L. A. Jr., KAISER J. F. — Analytical design of linear feedback controls. New York, John Wiley, 1957.

1.6. FELDBAUM A. — Critères intégraux de la régulation (en russe). « Automatique et télémécanique », 1948, t. 9, n° 1, pp. 3-19.

1.7. GIBSON G. — How to specify the performance of closed loop systems. Control Engineering, 1956, v. 3. Sept., pp. 122-129.

1.8. SHULTZ W. C., RIDEOUT V. C.— Performance measures: past, present and future. IRE Trans. Automat. Control, 1961, v. AC-6, Febr., n° 1, pp. 22-35.

1.9. FELDBAUM A.— Systèmes électriques de régulation automatique (en russe). Moscou, Oboronguiz, 1957.

1.10. POUGATCHEV V.— Théorie des fonctions aléatoires et son application aux problèmes de commande automatique (en russe). Moscou, Fizmatguiz, 1960.

1.11. LANING J. H., BATTIN R.— Random processes in automatic Control. New York, 1956.

1.12. DAVENPORT W. B., ROOT W. L.— An introduction to the theorie of random signals and noise. New York, Mc. Gram-Hill, 1958.

1.13. LÉVINE B.— Théorie des processus aléatoires et son application à la radiotechnique (en russe). Moscou, Sovetskoïé Radio, 1960.

1.14. KOLMOGOROV A.— Interpolation et extrapolation des séquences stationnaires aléatoires (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., série mathématique, 1941, t. 5, n° 1, pp. 1-14.

1.15. WIENER N.— Extrapolation, interpolation and smoothing of stationary time series. New York, John Wiley, 1949.

1.16. ZADEN L. A., RAGAZZINI T. R.— Extension of Wiener's theory of prediction. J. Appl. Phys., 1950, v. 21, n° 7, pp. 645-655.

1.17. ZADEN L. A.— Optimum non-linear filters. J. Appl. Phys. 1953, v. 24, pp. 396-404.

1.18. LUBBOCK J. K.— The optimization of a class of non-linear filters. IEE Proc., 1959, v. 107C, March, n° 11.

1.19. KHINTCHINE A. — Théorie de la corrélation des processus stochastiques stationnaires (en russe). « Succès des mathématiques », 1938, fasc. 5, pp. 42-51.

1.20. SEIFERT W. W., STEEG C. W. (ed).— Control systems engineering. New York. Mc Graw-Hill, 1960, p. 944.

1.21. LEE Y. N.— Statistical theory of communication. New York, John Wiley, 1960.

1.22. CHANG S. S. L.— Synthesis of optimum control systems. New York. Mc. Graw-Hill, 1961.

1.23. SOLODOV A. — Systèmes asservis linéaires aux paramètres variables (en russe). Moscou, Fizmatguiz, 1962.

1.24. WESTCOTT J. H.— Synthèse des systèmes asservis optimaux de puissance. Recueil des traductions. Moscou, 1955, n° 8.

1.25. PÉLEGRIN M.— Calcul statistique des systèmes asservis. Paris, Publications scientifiques et techniques du Ministère de l'Air, 1953.

1.26. ANDRÉEV N.— Condition générale d'extremum de la fonction donnée de l'écart-type et du carré de l'espérance mathématique de l'erreur d'un système dynamique (en russe). « Automatique et télémécanique », 1959, t. 30, n° 7, pp. 833-838.

1.27. WESTCOTT J. H.— Design of multivariable optimum filters. Trans. ASME, 1958, v. 80, p. 463.

1.28. KALMAN R. E., BUCY R. S.— New results in linear filtering and prediction theory. Trans. ASME, 1961, p. 95.

Chapitre II

2.1. GNÉDENKO B.— Cours de théorie des probabilités (en russe). Moscou, Fizmatguiz, 1961.

2.2. VENTSEL E.— Théorie des probabilités (en russe). Moscou, Fizmatguiz, 1962.

2.3. FELLER W.— An introduction to Probability theory and its application. New York, John Wiley, 1957.

2.4. SMIRNOV N., DOUNINE-BARKOVSKI I.— Cours abrégé de statistique mathématique pour des applications techniques (en russe). Moscou, Fizmatguiz, 1959.

2.5. SMIRNOV V., KRYLOV V., KANTOROVITCH L.— Calcul des variations (en russe). Kouboutch, 1933.

2.6. LAVRENTIEV M., LUSTERNIK L.— Cours de calcul des variations (en russe). Moscou-Léningrad, Gostechteoretizdat, 1950.

2.7. ELSGOLTZ L.— Calcul des variations (en russe). Moscou-Léningrad. Gostechteoretizdat, 1952.

2.8. GUELFAND I., FOMINE S.— Calcul des variations (en russe). Moscou, Fizmatguiz, 1961.

2.9. BELLMAN R.— Dynamic programming. Princeton, Princeton University Press, 1957.

2.10. BELLMAN R., DREYFUS S.— Functional approximations and dynamic programming. Mathematical tables and other aids to computation. 1959, v. 13, Oct., n° 68.

2.11. BOLTIANSKI V., GAMKRÉLIDZÉ R., PONTRIAGUINE L.— Contribution à la théorie des processus optimaux (en russe). Comptes-rendus de l'Académie des Sciences de l'U.R.S.S., 1956, t. 110, n° 1, pp. 7-10.

2.12. GAMKRÉLIDZÉ R.— Théorie des processus en temps minimal dans des systèmes linéaires (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., série mathématique, 1958, t. 22, n° 4, pp. 449-474.

2.13. BOLTIANSKI V.— Principe du maximum dans la théorie des processus optimaux (en russe). « Doklady » de l'Académie des Sciences de l'U.R.S.S. 1958, t. 119, n° 6. pp. 1070-1073.

2.14. GAMKRÉLIDZÉ R.— Contribution à la théorie générale des processus optimaux (en russe). « Doklady » de l'Académie des Sciences de l'U.R.S.S., 1958, t. 123, n° 2, pp. 223-226.

2.15. ROSONOER L.— Sur les conditions d'optimalité suffisantes (en russe). « Doklady » de l'Académie des Sciences de l'U.R.S.S., 1959, t. 127, n° 3, pp. 520-523.

2.16. PONTRIAGUINE L.— Processus optimaux de régulation (en russe). « Succès des mathématiques », 1959, t. 14, fasc. 1, pp. 3-20.

2.17. ROSONOER L.— Principe du maximum de Pontriaguine L. dans la théorie des systèmes optimaux (en russe). « Automatique et télémécanique », 1959, t. 20, n° 10, pp. 1320-1334; n° 11, pp. 1441-1458; n° 12, pp. 1561-1578.

2.18. GAMKRÉLIDZÉ R.— Processus en temps minimal à coordonnées de phase limitées (en russe). « Doklady » de l'Académie des Sciences de l'U.R.S.S., 1959, t. 125, n° 3, pp. 457-478.

2.19. ROSONOER L.— Application des méthodes variationnelles à l'étude des performances des systèmes asservis (en russe). Moscou, Editions de l'Académie des Sciences de l'U.R.S.S., 1960.

2.20. DESOER C. A.— Pontriagin's maximum principle and the principle of optimality. J. Franklin Inst., 1961, v. 271, n° 5, pp. 361-367.

2.21. PONTRIAGUINE L., BOLTIANSKI V., GAMKRÉLIDZÉ R., MITCHTCHENKO E.— Théorie mathématique des processus optimaux (en russe). Moscou, Fizmatguiz, 1961.

2.22. KRASSOVSKI N.— Contribution à la théorie de commande optimale (en russe). « Automatique et télémécanique », 1957, t. 18, n° 11, pp. 960-970.

2.23. KRASSOVSKI N.— Sur un problème de commande optimale (en russe). « Mathématiques appliquées et mécanique », 1957, t. 21, n° 5, pp. 670-677.

2.24. AKHIÉZER N., KREINE M.— Sur certaines questions de la théorie des moments (en russe). Kharkov, GONTI, 1938.

2.25. KULIKOWSKI R.— Concerning a class of optimum control systems. Bull. Academ. Polon. Scient., 1960, v. 8, n° 10, pp. 595-600.

2.26. KULIKOWSKI R.— O ukladach optymalnych regulacji automatycznej (en polonais). Archiwum Automat. i Telemech., 1961, t. 6, n° 2-3, p.p. 236-296.

2.27. KIRILLOVA F.— Certaines questions de la théorie de régulation optimale (en russe). « Izvestia » des écoles supérieures. Mathématique, 1962, n° 3, pp. 48-58.

2.28. GAMKRÉLIDZÉ R.— Sur les régimes optimaux glissants (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., 1962, t. 143, n° 6, pp. 1243-1245.

2.29. KROTOV V.— Méthodes de résolution des problèmes variationnels basées sur les conditions suffisantes du minimum absolu, partie I (en russe). « Automatique et télémécanique », 1962, t. 23, n° 12, pp. 1571-1583.

2.30. SCHREIDER Y.— Problèmes de programmation dynamique et automates (en russe). Recueil « Problèmes de cybernétique », Moscou, 1961, fasc. 5, pp. 31-48.

2.31. BOLTIANSKI V.— Conditions d'optimalité suffisantes (en russe). Comptes-rendus de l'Académie des Sciences de l'U.R.S.S., 1961, t. 140, n° 5, pp. 994-997.

2.32. MIELE A.— General variational theory of the flight paths of rocket powered aircraft missiles and satellite carriers. Astronaut. Acta, 1958, v. 4, n° 4, pp. 264-288.

2.33. BERKOVITZ L. D.— Variational methods in problems of control and programming. J. Mathem. Analys. and Applic., 1961, v. 3. Aug. n° 1, pp. 145-169.

2.34. TROÏTSKI V.— Problèmes variationnels d'optimisation de la commande dans les systèmes à coordonnées limitées (en russe). « Mathématiques appliquées et mécanique », 1962, t. 26, fasc. 3, pp. 431-443.

2.35. BELLMAN R., DREYFUS S.— Applied dynamic programming. Princeton, Princeton University Press, 1962.

2.36. DESOER C. A.— The bang-bang servo problem treated by variational techniques. Information and control, 1959, v. 2, n° 4, pp. 333-348.

2.37. CHANG S.— Contribution à la théorie de commande optimale (en russe). « Mathématiques appliquées et mécanique », 1961, t. 25, fasc. 3, pp. 413-419.

2.38. BELLMAN R.— Adaptive control processes: a guided tour, New Jersey, Princeton University Press, 1961.

2.39. CHANG S. S. L.— Minimal time control with multiple saturation limits. IRE International Convention Record, 1962, pp. 143-151.

2.40. KRYLOV N. et TCHERNOOUSKO F.— Application de la méthode des approximations successives à la résolution des problèmes de commande optimale (en russe). « Revue de calcul numérique et de physique mathématique », 1962, t. 2, n° 6, pp. 1132-1139.

2.41. KATZ S.— A discrete version of Pontriagin's maximum principle. J. Electron. and Control, Ser. 1, 1962, v. 13, Aug., n° 2, p. 179-184.

2.42. GABASSOV R.— Contribution au problème de l'unicité de la commande optimale dans des systèmes discrets (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., OTN, Energétique et automatique, 1962, n° 5, pp. 99-106.

2.43. PROPOÏ A.— Sur le principe du maximum dans des systèmes de comande discrets (en russe). « Automatique et télémécanique », n° 7, 1965.

2.44. LEITMANN G. (ed.) — Optimization techniques with applications to aerospace systems. Academic Press, New York-London, 1962.

2.45. CHANG S.S.L.— Optimal control in bounded phase space. Automatica, January-March 1963, pp. 55-67.

2.46. CHANG S. S. L.— Minimal time control with multiple saturation limits. IEEE Transactions on Autom. Control, January 1963, n° 1, v. AC-8, pp. 35-42.

2.47. BOUTKOVSKI A.— Principe du maximum étendu aux problèmes de commande optimale (en russe). « Automatique et télémécanique », t. XXIV, n° 3, pp. 314-327.

2.48 DOUBOVITSKI A., MILUTINE A.— Problèmes d'extremum en présence des contraintes (en russe). « Doklady » de l'Académie des Sciences, 1963, t. 149, n° 4, p. 759.

Chapitre III

3.1. MARIANOVSKI D., SVETCHARNIK D.— Brevet n° 77023, demande 181007 du 25 février 1935.

3.2. FELDBAUM A.— Systèmes asservis à relais élémentaires (en russe). « Automatique et télémécanique », 1949, t. 10, n° 4, pp. 249-266.

3.3. HOPKIN A. M.— A phase-plane approach to the compensation of saturating servomechanisms. Trans. AIEE, 1951, Pt. 1, v. 70, pp. 631-639.

3.4. LERNER A.— Amélioration des performances dynamiques des compenseurs automatiques à l'aide des non-linéarités (en russe). « Automatique et télémécanique », 1952, t. 13, n° 2, pp. 134-144; n° 4, pp. 429-444.

3.5. FELDBAUM A.— Processus optimaux dans les systèmes asservis (en russe). « Automatique et télémécanique », 1953, t. 14, n° 6, pp. 712-728.

3.6. LERNER A.— Sur la commande en temps minimal des systèmes asservis (en russe). « Automatique et télémécanique », 1954, t. 15, n° 6, pp. 461-477.

3.7. BOGNER I., KAZDA L. F.— An investigation of the switching criteria for higher order contactor servomechanisms. Trans. AIEE, 1954, Pt. 2, v. 73, pp. 118-127.

3.8. SILVA L. M.— Predictor servomechanisms. Trans. IRE, 1954, v. CT-1, March, n° 1, pp. 36-70.

3.9. LA SALLE J. P.— Basic principle of the bang-bang servo. Bull. Amer. Math. Soc., 1954, v. 60, n° 2, p. 154.

3.10. FELDBAUM A.— Contribution à la synthèse des systèmes asservis optimaux (en russe). Travaux de la IIe Conférence nationale sur la théorie de la régulation automatique (1953). Moscou, Académie des Sciences de l'U.R.S.S., 1955, pp. 325-360.

3.11. FELDBAUM A.— Sur la synthèse des systèmes optimaux par la méthode d'espace de phase (en russe). « Automatique et télémécanique », 1955, t. 16, n° 2, pp. 120-149.

3.12. HOPKIN A. M., IWAMA M.— A study of a predictor type « Air-Frame » controller designed by phase-space analysis. Application and Industry, 1956, March, n° 23, pp. 1-9.

3.13. KUBA R. E., KAZDA L. F.— A phase-space method for the synthesis of non-linear servomechanisms. Application and Industry, 1956, Nov., n" 27, pp. 282-289.

3.14. BASS R. W.— Equivalent linearization, non-linear circuit synthesis and the stabilization and optimization of control system. Proc. of the Symposium on non-linear circuit analysis. Polytechn. Inst. of Brooklin, April, 1956.

3.15. TSIANG SU-SÈN — Cybernétique technique.

3.16. ROSENMAN E.— Sur les processus transitoires optimaux dans des systèmes à puissance limitée (en russe). « Automatique et télémécanique », 1957, t. 18, n° 6, pp. 497-513.

3.17. BUSHAW D. W.— Optimal discontinuous forcing terms. Contribution to non-linear oscillation. Annals of mathematical study, v. 41. Princeton, Princeton University Press, 1958.

3.18. LA SALLE J. P.— Time optimal control systems. Proc. of National Acad. Sci., 1959, v. 45.

3.19. NOVOSELTSEV V.— Processus optimal dans un système à relais impulsionnel du second ordre (en russe). « Automatique et télémécanique », 1960, t. 21, n° 5, pp. 569-574.

3.20. DESOER C. A., WING J.— An optimal strategy for a saturating sampled data system. IRE Trans. Automat. Control, 1961, v. AC-6, Febr., n° 1, pp. 5-15.

3.21. SMITH F. B.— Time-optimal control of higher order systems. IRE Trans. Automat. Control, 1961, v. AC-6, Febr., n° 1.

3.22. GOSIEWSKI A.— Optymalne servomechanismy przekaznikowe drugiego rzedu (en polonais). Polska Akademia Nauk, Rozprawy Elektrotechniczne, 1961, t. 7, n° 1, pp. 17-67.

3.23. CHANDAKET P., LEONDES C. T.— Optimum non-linear bang-bang control systems with complex roots. Applications and Industry, 1961, May.

3.24. DESOER C. A., WING J.— A minimal time discrete system. IRE Trans. Automat. Control, 1961, v. AC-6, May, n° 2.

3.25. FELDBAUM A.— Calculateurs dans des systèmes asservis (en russe). Moscou, Fizmatguiz, 1959.

3.26. KULIKOWSKI R.— Concerning a class of optimum control systems. Bull. Acad. Polon. Sci., Ser. Sci. techn., 1960, v. 8, n° 10, pp. 595-600.

3.27. KIPINIAK W.— Dynamic optimization and control. A variational approach. Mass. Inst., 1961.

3.28. LÉTOV A.— Conception analytique des régulateurs (en russe). « Automatique et télémécanique », 1960, t. 21, n° 4, pp. 436-441 ; n° 5, pp. 561-568 ; n° 6, pp. 661-665.

3.29. OSTROVSKI G.— Sur une méthode de résolution des problèmes variationnels (en russe). « Automatique et télémécanique », 1962, t. 23, n° 10, pp. 1284-1289.

3.30. TSYPKINE I.— Sur les processus optimaux dans des systèmes asservis impulsionnels (en russe). « Doklady » de l'Académie des Sciences de l'U.R.S.S., 1960, t. 134, n° 2, pp. 308-310.

3.31. TSYPKINE I.— Processus optimaux dans des systèmes asservis impulsionnels (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., « Energétique et automatique », 1960, n° 4, pp. 74-93.

3.32. ECKMAN D. P., LEFKOWITZ J.— A report on ... optimizing control of a chemical process. Control Engng, 1957, n° 9, pp. 197-204.

3.33. GOULD L. A., KIPINIAK W.— Dynamic optimization and control of a stirred-tank chemical reactor. Communic. and Electronics, 1961, Jan., pp. 734-743.

3.34. MERRIAM C. W.— Synthesis of adaptive controls. Mass. Inst. of Technology, 1959, Febr.

3.35. MERRIAM C. W.— A class of optimum control systems. J. Franklin Inst., 1959, v. 267, n° 4, pp. 267-281.

3.36. PETERSON E. L.— Statistical analysis and optimization of systems. New York, John Wiley, 1961.

3.37. SUN TSIAN — Synthèse d'un système de commande optimal sur la base d'un champ d'isochrones (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., « Energétique et automatique», 1960, n° 5, pp. 96-103.

3.38. ZADEH L. A.— Optimal control problems in discrete-time systems. Dans: LEONDES T. (ed.). Computer control systems technology. New York, Mc Graw-Hill, 1961, pp. 389-414.

3.39. CHANG JEN-VEY — Un problème de synthèse des systèmes optimaux d'après le principe du maximum (en russe). « Automatique et télémécanique », 1961 t. 22, n° 10, pp. 1302-1308.

3.40. SUN TSIAN — Commande optimale d'un système non linéaire (en russe). « Automatique et télémécanique », 1960, t. 21, n° 1, pp. 3-14

3.41. BOR-RAMENSKI A., SUN TSIAN — Commande asservie optimale à deux paramètres de commande (en russe). « Automatique et télémécanique », 1961, t. 22, n° 2, pp. 157-170.

3.42. BOUTKOVSKI A., LERNER A.— Sur la commande optimale des systèmes à paramètres répartis (en russe). « Automatique et télémécanique », 1960, t. 21, n° 6, pp. 682-691.

3.43. BOUTKOVSKI A.— Processus optimaux dans des systèmes aux paramètres répartis (en russe). « Automatique et télémécanique ». 1961, t. 22, n° 1, pp. 17-26.

3.44. BOUTKOVSKI A.— Principe du maximum appliqué aux systèmes optimaux à paramètres répartis (en russe). « Automatique et télémécanique », 1961, t. 22, n° 10, pp. 1288-1301.

3.45. BELLMAN R.— On the application of the theory of dynamic programming to the study of control processes. Proc. of the Symposium on nonlinear circuit analysis. Polytechn. Inst. of Brooklin, April, 1956, pp. 199-213.

3.46. BELLMAN R., GLICKSBERG I., GROSS O.— On the bang-bang control problem. Quarterly of Applied Mathematics, 1956, v. 14.

3.47. KALMAN R. E., KOEPCKE R. W.— Optimal synthesis of linear sampling control systems using generalized performance indexes. Trans. ASME, 1958, v. 80, pp. 1820-1826.

3.48. JOHNSON C. D., WONHAM W. M.— On a problem of Letov in optimal control. Trans. ASME, v. 85, ser. D, n° 2, 1965.

3.49. KRAMER J. D. H.— On control of linear systems with time lags, Information and Control, 1960, v. 3, pp. 299-326.

3.50. SMITH F. B.— Time optimal control of higher-order systems. IRE, Trans, Automat. Control, 1961, v. AC-6, Febr., n° 1, pp. 16-21.

3.51. CHANG S. S. L.— Computer optimization of non-linear control systems by means of digitized maximum principle. IRE International Convention Record, 1961, Pt. 4, pp. 48-55.

3.52. ARONOVITCH G.— Sur l'influence des phénomènes ondulatoires dans une conduite à pression sur la commande optimale (en russe). « Izvestia » des écoles supérieures. Radiophysique, 1962, t. 5, n° 2, pp. 362-369.

3.53. KRASSOVSKI N., LETOV A.— Contribution à la théorie de conception analytique des régulateurs (en russe). « Automatique et télémécanique », 1962, t. 23, n° 6, pp. 713-720.

3.54. GABASSOV R.— Sur les processus optimaux dans des systèmes discrets liés (en russe). « Automatique et télémécanique », 1962, t. 23, n° 7, pp. 872-880.

3.55. BALAKIREV V.— Principe du maximum dans la théorie des systèmes optimaux de second ordre (en russe). « Automatique et télémécanique », 1962, t. 23, n° 8, pp. 1014-1022.

3.56. KRANC G. M., SARACHIK P. E.— An application of functional analysis to the optimal control problem. Dans: Joint automatic control conference (ASME, ISA, AIEE, AJChE), June 27-29, 1962, 8-2, pp. 1-8.

3.57. KAZDA L. F.— Application of state variables to optimal control systems problems. Dans: Joint automatic control conference (ASME, ISA, AIEE, IRE, AJChE), 1962, 11-2, pp. 1-5.

3.58. FLUGGE-LOTZ I., MARBACH H.— The optimal control of some attitude control systems for different performance criteria. Dans: Joint automatic control conference (ASME, ISA, AIEE, AJChE), 1962, 12-1, pp. 1-12.

3.59. NEUSTADT L.— Time optimal control systems with position and integral limits. J. Mathem. Analysis and Applic., 1961, v. 3, Dec., n° 3.

3.60. FRIEDLAND B.— The structure of optimum control systems. Trans. ASME Journ. of Basic Engng, Ser. D., 1962, v. 84, n° 1, pp. 1-13.

3.61. ISSAÉV V., SONINE V.— Sur un problème non linéaire de commande optimale (en russe). « Automatique et télémécanique », 1962, t. 23, n° 9, pp. 1117-1129.

3.62. NHUEN THAN BANG — Sur la résolution de certains problèmes de la théorie de programmation dynamique à l'aide des calculateurs analogiques (en russe). « Automatique et télémécanique », 1962, t. 23, n° 9, pp. 1130-1140.

3.63. PARAÉV I.— Sur une commande spéciale des processus optimaux linéaires par rapport aux actions de commande (en russe). « Automatique et télémécanique », 1962, t. 23, n° 9, pp. 1202-1209.

3.64. SAVVINE A.— Sur la théorie des processus en temps minimal dans des systèmes de second ordre (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., « Energétique et automatique », 1960, n° 6, pp. 162-164.

3.65. LITVINE-SÉDOÏ M., SAVVINE A.— Sur la synthèse des systèmes asservis de second ordre aux processus transitoires limités (en russe). « Vestnik » de l'Université de Moscou, sér., 1, Math. et méc., 1961, t. 4, n° 5, pp. 77-83.

3.66. BRENNAN P. J., ROBERTS A. P.— Use of an analogue computer in the application of Pontriagin's maximum principle to the design of control systems with optimum transient response. J. Electron. and Control, Ser. 1, 1962, v. 12, n° 4, pp. 345-352.

3.67. PAVLOV A.— Sur la diminution du temps de réponse de certains systèmes à relais de troisième ordre (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., « Energétique et automatique », 1962, n° 2, pp. 59-71.

3.68. ATHANASSIADES M., SMITH O. J. M.— Theory and design of high-order bang-bang control systems. IRE Trans. Automat. Control, 1961, v. AC-6, May, pp. 125-134.

3.69. FRIEDLAND B.— A minimum response-time controller to amplitude and energy constraints. IRE Trans. Automat. Control, 1962, v. AC-7, Jan., n° 1, pp. 73-74.

3.70. NELSON W. L.— Optimal control methods for on-off sampling systems. Trans. ASME Basic Engng, 1962, March, pp. 91-99.

3.71. FELDBAUM A. — Sur la synthèse automatique des systèmes asservis optimaux. Calculateurs combinés (en russe). Travaux de la Conférence sur la théorie et l'histoire de la simulation mathématique, Moscou, Editions de l'Académie des Sciences de l'U.R.S.S., 1962.

3.72. KALLAY M.— An exemple of the application of dynamic programming to the design of optimal control programms. IRE Trans. Automat. Control, 1962, Apr., n° 3, p. 10.

3.73. WAZEWSKI T.— O problemie optymalnego sterowania w przypadku nieliniowym (en polonais). Arch., Automat. i Telemech., 1962, t. 7, n°s 1-2, pp. 19-32.

3.74. KIRILLOVA M.— Un problème d'optimisation de l'état final d'un système commandé (en russe). « Automatique et télémécanique », 1962, t. 23, n° 12, pp. 1584-1594.

3.75. SALOUKVADZE M.— Sur la synthèse d'un régulateur optimal dans des systèmes linéaires à retard soumis à des perturbations permanentes (en russe). « Automatique et télémécanique », 1962, t. 23, n° 12, pp. 1595-1601.

3.76. OUTKINE V.— Construction d'une classe de systèmes asservis sans utiliser dans la loi de commande de dérivées « pures » non perturbées (en russe). « Automatique et télémécanique », 1962, t. 23, n° 12, pp. 1631-1642.

3.77. LERNER A.— Principe de conception des systèmes asservis et régulateurs en temps minimal (en russe). Moscou-Léningrad, Gosénergoïzdat, 1961.

3.78. JOHNSON C. D., GIBSON J. E.— Singular solutions in problems of optimal control. IEEE Trans. Automat. Control, January 1963, n° 1, v. AC-8, pp. 4-15.

Chapitre IV

4.1. ROÏTENBERG I.— Sur certaines méthodes indirectes d'obtention de l'information sur l'état du système gouverné dans un espace de phase (en russe). « Math. appl. et méc. », 1961, t. 25, fasc. 3, pp. 440-444.

4.2. FULLER A. T.— Optimization of non-linear control systems with random inputs. J. Electron. and Control., Ser. 1, 1960, v. 9, July, n° 1, pp. 65-80.

4.3. FLORENTIN J. J.— Optimal control of continuous time Markov sthocastic systems. J. Electron. and Control, Ser. 1, 1961, v. 10, n° 6.

4.4. AOKI M.— Stochastic-time optimal control systems. Applications and Industry, 1961, May, n° 54, pp. 41-44; Disc., pp. 44-46.

4.5. KRASSOVSKI N., LIDSKI E.— Conception analytique des régulateurs dans des systèmes aux propriétés aléatoires (en russe). « Automatique et télémécanique », 1961, t. 22, n° 9, pp. 1145-1150; n° 10, pp. 1273-1278; n° 11, pp. 1425-1431.

4.6. NOVOSELTSEV V.— Commande optimale d'un système à relais impulsionnel de second ordre en présence des perturbations aléatoires (en russe). « Automatique et télémécanique », 1961, t. 22, n° 7, pp. 865-875.

4.7. HOWARD A. R.— Dynamic programming and Markov processes. Technology Press of MIT, 1960.

4.8. BELLMAN R.— A Markovian decision process. J. Mathem. and Mechan., 1957, v. 6, p. 679.

4.9. HIGGINS W. H.— Signal-flow graphs and random signals. P.I.R.E., 1957, v. 45, p. 74.

4.10. SITTLER P. W.— System analysis of discrete Markov processes. Trans. IRE, PGCT, 1956, v. CT-3, n° 4.

4.11. FELDBAUM A.— Processus stable dans un système discret extrémal simple en présence des perturbations aléatoires (en russe). « Automatique et télémécanique », 1959, t. 20, n° 8, pp. 1056-1070.

4.12. EATON J. H., ZADEH L. A.— Optimal pursuit strategies in discrete state probabilistic systems. Trans. ASME, J. of Basic Engng, 1962, March, pp. 23-29.

4.13. ROMANOVSKI V.— Chaînes discrètes de Markov (en russe). Moscou, Gostekhizdat, 1947.

4.14. FELLER W.— An introduction in probability theory and its applications. New York, John Wiley, 1957.

4.15. GUELFOND A.— Calcul des différences finies. Paris, Dunod, 1966.

4.16. TSYPKINE I.— Théorie des systèmes impulsionnels. Moscou, Fizmatguiz, 1958.

4.17. BUSH R. R., MOSTELLER F.— Stochastic models for learning. New York, John Wiley, 1955.

4.18. BUSH R. R., ESTES W. (ed.) — Studies in mathematical learning theory. Standford University, 1959.

4.19. NOVOSELTSEV V.— Systèmes de commande en temps minimal en présence des perturbations aléatoires (en russe). « Automatique et télémécanique », 1962, t. 23, n° 12, pp. 1620-1630.

4.20. BELLMAN R.— Dynamic programming and stochastic control processes. Information and Control, 1958, Sept., n° 3.

4.21. FLORENTIN J. J.— Optimal control of systems with generalized Poisson inputs. Dans le recueil: Joint automatic control conference (ASME, ISA, AIEE, IRE, AIChE), June 27, 28, 29, 1962, 14-2, pp. 1-5.

4.22. KATZ S.— Best endpoint control of noisy systems. J. Electron. and Control, 1962, v. 12, n° 4, pp. 323-343.

4.23. ANDRONOV A., VITT A., PONTRIAGUINE L.— Sur l'étude statistique des systèmes dynamiques (en russe). Journ. de physique technique expérimentale, 1933, fasc. 3, pp. 165-180.

4.24. BARRETT J. F.— Application des équations de Kolmogorov à l'étude des systèmes asservis aux perturbations aléatoires (en russe). Dans le recueil: Travaux du Premier Congrès International de l'IFAC sur la commande automatique, 1960, Moscou, Editions de l'Académie des Sciences de l'U.R.S.S., 1961, t. 3, pp. 84-105.

4.25. CHEN P. P.— An algorithm for stochastic control through dynamic programming techniques. IRE Intern. Conven. Record, 1962, Pt. 2, pp. 112-124.

4.26. STRATONOVITCH R., SCHMALHAUSEN V.— Certains problèmes stationnaires de programmation dynamique (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., Energétique et automatique, 1962, n° 5, pp. 131-139.

4.27. CHEN P. P.— An algorithm for stochastic control through dynamic programming. IRE Trans. Autom. Control, v. AC-7, Oct. 1962, n° 5, pp. 93-101.

Chapitre V

5.1. KOULÉBAKINE V.— Asservissements invariants de haute qualité. Théorie de l'invariance et son application à l'automatique (en russe)., Travaux de la Conférence de Kiev du 16 au 20 octobre 1958, Moscou Académie des Sciences de la R.S.S. d'Ukraine, 1959, p. 11-39.

5.2. PÉTROV B.— Principe d'invariance] et conditions de son application au calcul des systèmes linéaires et non linéaires (en russe). Travaux du premier Congrès International de l'IFAC sur la commande automatique. Moscou, Editions de l'Académie des Sciences de l'U.R.S.S., 1961, t. 1, pp. 159-275.

5.3. SMITH O. J. M.— Feedback control systems. Mc Graw-Hill, New York, 1958.

5.4. MÉÉROV M.— Synthèse des structures des systèmes asservis de haute précision (en russe). Moscou, Fizmatguiz, 1959.

5.5. KOGAN B.— Calculateurs analogiques et leur application à l étude des systèmes asservis (en russe). Moscou, Fizmatguiz, 1959.

5.6. CHTCHIPANOV G.— Théorie et méthodes d'établissement des projets de régulateurs automatiques (en russe). « Automatique et télémécanique », 1939, n° 1, pp. 49-66.

5.7. CRAMER H.— Mathematical methods of statistics. Princeton, Princeton Univ. Press, 1946.

5.8. SMIRNOV N., DOUNINE-BARKOVSKI I.— Cours abrégé de statistique mathématique pour des applications techniques (en russe). Moscou, Fizmatguiz, 1959.

5.9. LINNIK I.— Méthode des moindres carrés et principes de la théorie statistique du traitement des observations (en russe). Moscou, Fizmatguiz, 1962.

5.10. WILLIAMS D. D.— The compleat strategyst being a primer on the theory of games of strategy. New York, Mc Graw-Hill, 1954.

5.11. VENTSEL E.— Eléments de la théorie des jeux (en russe). Moscou, Fizmatguiz, 1959.

5.12. VAJDA S.— Théorie des jeux et programmation linéaire. Dunod, Paris, 1959.

5.13. Mac KINSEY J. C. C.— An introduction to theory of games. New York, Mc Graw-Hill, 1952.

5.14. LUCE R. D., RAIFFA H.— Games and decisions. New York, John Wiley, 1957.

5.15. WALD A.— tatistical decision functions. New York, John Wiley, 1950.

5.16. KOTELNIKOV V.— Théorie de la stabilité aux perturbations potentielle en présence des bruits de fluctuation (en russe). Thèse de doctorat, Institut énergétique de Moscou, Moscou, 1946.

5.17. VAÏNSTEIN L., ZOUBAKOV V.— Sélection des signaux en présence des perturbations aléatoires (en russe). Moscou, Sovetskoïe radio, 1960.

5.18. GOUTKINE L. (sous la direction de) — Réception des signaux en présence des bruits. Recueil de traductions (en russe). Moscou, Editions de la littérature en langues étrangères, 1960.

5.19. BACHARINOV A., ALEXANDROV M (sous la direction de) — Réception des signaux impulsionnels en présence des bruits. Recueil des traductions (en russe). Moscou-Léningrad, Gosénergoïzdat, 1960.

5.20. FALKOVITCH S.— Réception des signaux radar en présence des bruits de fluctuation (en russe). Moscou, Sovetskoïe radio, 1961.

5.21. GOUTKINE A.— Théorie des méthodes optimales de la réception radio en présence des bruits de fluctuation (en russe). Moscou-Léningrad, Gosénergoïzdat, 1961.

5.22. MIDDELTON D.— An introduction to statistical communication theory, t. 1-2, 1960.

5.23. CHOW C. K.— An optimum character recognition system using decision functions. IRE Trans. on Electronic Computers, 1957, Dec., v. EC-6, n° 4, pp. 247-254.

5.24. FELDBAUM A.— Théorie de la commande duale, II (en russe). « Automatique et télémécanique », 1960, t. 21, n° 11, pp. 1453-1464.

5.25. TRUXALL J. G., PADALINO J. J.— Decision theory. Dans le recueil: Adaptive Control Systems, ed. by Eli Mishkin and Ludwig Braun J. R., Mc Graw-Hill, New York, 1961.

5.26. RYJIK I.— Tables des intégrales, sommes, séries et produits. Moscou, Gostekhizdat, 1948.

5.27. TCHERNOV G., MOZES L.— Théorie élémentaire des décisions statistiques (en russe). Sovetskoïe radio, 1952.

5.28. TARASSENKO V.— Sur les méthodes optimales de régulation extrémale en présence des perturbations (en russe). Dans le recueil: Calculateurs. Automatique. Informatique. Université de Tomsk, 1961, fasc. 40, pp. 47-57.

5.29. HSU J. C., MESERVE W. E.— Decision-making in adaptive control systems. IRE Trans. Automat. Control. 1962, v. AC-7, n° 1, pp. 24-32.

5.30. VALD A.— Echantillonnage progressif (en russe). Moscou, Fizmatguiz, 1960.

5.31. BACHARINOV A., FLEISCHMAN B.— Méthodes statistiques de l'échantillonnage progressif et leurs applications radiotechniques (en russe). Moscou, Sovetskoïe radio, 1962.

5.32. BELLMAN R., KALABA R.— On adaptive control processes. IRE Nation. Convention Record, Pt. 4, 1959, July, v. 7, pp. 3-11.

5.33. FREIMER M. A.— Dynamic programming approach to adaptive control processes. IRE Nation. Convention Record, Pt. 4, 1959, v. 7, n° 4, pp. 12-17.

5.34. KÉLENDJÉRIDZE D.— Sur un problème de poursuite optimale (en russe). « Automatique et télémécanique », 1962, t. 23, n° 8, pp. 1008-1013.

5.35. GADJIEV M.— Application de la théorie des jeux à certains problèmes de commande automatique (en russe). « Automatique et télémécanique », 1962, t. 23, n° 8, pp. 1023-1036; n° 9, pp. 1144-1153.

5.36. RÉPINE V.— Sur la résolution d'un problème de filtrage dans les conditions des statistiques a priori incomplètes (en russe). « Automatique et télémécanique », 1962, t. 23, n° 8, pp. 1108-1111.

5.37. CARNAP R.— Logical foundations of probability, 1950.

5.38. SCHWARTZ L. S., HARRIS B., HAUPTSCHEIN A.— Information rate from the viewpoint of inductive probability. IRE Nation. Convention Record, Pt. 4, 1959.

5.39. KIEFER J., WOLFOWITZ J.— Optimum design in regression problems. Ann. Math. Statist., 1959, v. 30, pp. 271-294.

5.40. BELLMAN R., KALABA R.— Dynamic programming and modern control theory. New York, Acad. Press, 1965.

5.41. KUSHNER H. J.— Optimal stochastic control. URE Trans. Automat. Control, Oct. 1962, v. AC-7, n° 5, pp. 120-122.

5.42. ROSONOER L.— Approche variationnelle du problème d'invariance (en russe). « Automatique et télémécanique », 1963, t. 24, n° 6, pp. 744-752; n° 7, pp. 861-870.

Chapitre VI

6.1. FELDBAUM A.— Théorie de la commande duale (en russe). « Automatique et télémécanique », 1960, t. 21, n° 9, pp. 1240-1249; n° 11, pp. 1453-1464; 1961, t. 22, n° 1, pp. 3-16; n° 2, pp. 129-142.

6.2. KAZAKÉVITCH V.— Systèmes de régulation extrémale et certains procédés d'amélioration de leur stabilité (en russe). Dans le recueil : Commande automatique et calcul numérique. Moscou, Machguiz, 1958, pp. 66-96.

6.3. MOROSSANOV I.— Méthodes de régulation extrémale (en russe). « Automatique et télémécanique », 1957, t. 18, n° 11, pp. 1029-1044.

6.4. FELDBAUM A.— Sur l'utilisation des calculateurs électroniques dans des systèmes automatiques (en russe). « Automatique et télémécanique », 1956, t. 17, n° 11, pp. 1046-1056.

6.5. FITSNER L.— Sur les principes de construction et les méthodes d'étude de certains types de systèmes extrémaux (en russe). Travaux de la Conférence de 1958 sur la théorie et l'application des systèmes automatiques discrets, Moscou, Editions de l'Académie des Sciences de l'U.R.S.S., 1960, pp. 480-504.

6.6. FELDBAUM A.— Optimiseur automatique (en russe). « Automatique et télémécanique », 1958, t. 19, n° 8, pp. 731-743.

6.7. FELDBAUM A.— Sur l'accumulation de l'information dans des systèmes asservis en boucle fermée (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., Energétique et automatique, 1961, n° 4, pp. 107-119.

6.8. VAN DER WAERDEN B. L.— Matematische statistik. Berlin, Springer-Veil., 1957.

6.9. CRAMER H.— Mathematical methods of statistics. Princeton, Princeton Univ. Press, 1958.

6.10. STRATONOVITCH R.— Contribution à la théorie de la commande optimale. Coordonnées suffisantes (en russe). « Automatique et télémécanique », 1962, t. 23, n° 7, pp. 910-917.

6.11. FELDBAUM A.— Sur la commande optimale des systèmes markoviens (en russe). « Automatique et télémécanique », 1962, t. 23, n° 8, pp. 993-1007.

6.12. FELDBAUM A.— Sur les problèmes de la théorie de la commande duale. Rapport présenté au 2e Congrès international de l'IFAC (Bâle), 1963.

6.13. PIKE E. W., SILVERBERG T. R.— Designing mechanical computers. Machine Design, 1952, v. 24, n° 7, pp. 131-137; n° 8, pp. 159-163.

6.14. ETERMAN I.— Calculateurs analogiques (en russe). Moscou, Machguiz, 1957.

6.15. JIGOULEV V.— Synthèse des systèmes optimaux d'une classe (en russe). « Automatique et télémécanique », 1962, t. 23, n° 11, pp. 1431-1438.

6.16. FLORENTIN J. J.— Optimal, probing, adaptive control of a simple Bayesian system. J. Electron. and Control, First Series, 1962, v. 13, Aug., n° 2, pp. 165-177.

6.17. FELDBAUM A.— Sur l'application de la théorie des décisions statistiques aux systèmes asservis en chaîne ouverte et en boucle fermée (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., sér. « Cybernétique technique », 1963, n° 1.

6.18. FLORENTIN J. J.— Partial observability and optimal control. J. Electron. and Control. Ser. 1, Sept. 1962, v. 13, n° 3, pp. 263-279.

6.19. FULLER A. T.— Bibliography of optimum non-linear control of determinate and stochastic-definite systems. J. Electron. and Control. First Series, Dec. 1962, v. 13, n° 6, p. 588.

Chapitre VII

7.1. ROUTMAN R.— Systèmes auto-adaptatifs avec adaptation d'après les caractéristiques dynamiques (aperçu en russe). « Automatique et télémecanique », t. 23, n° 5, 1962.

7.2. SYSSOEV L.— Détection d'un signal aléatoire reçu par plusieurs canaux en présence de bruits indépendants (en russe). « Automatique et télémécanique », t. 24, n° 5, 1963.

7.3. KLOVSKI D.— Stabilité potentielle aux perturbations lors d'une réception en diversité d'information discrète (en russe). « Radiotechnique », n° 3, 1961.

7.4. MASLOV E.— Application de la théorie des décisions statistiques aux problèmes d'estimation des paramètres d'un système (en russe). « Automatique et télémécanique », t. 24, n° 10, 1963.

7.5. JIVOGLIADOV V., MASLOV E.— Sur un problème d'estimation des paramètres essentiels d'un système (en russe). « Izvestia » de l'Académie des Sciences de l'U.R.S.S., « Cybernétique technique », n° 1, 1964.

7.6. MASLOV E.— Estimation des paramètres des systèmes markoviens (en russe). « Automatique et télémécanique », t. 25, n° 1, 1964.

7.7. MASLOV E.— Modèle statistique auto-adaptatif I (en russe). « Automatique et télémécanique », t. 25, n° 10, 1964.

7.8. MASLOV E.— Modèle statistique auto-adaptatif II (en russe). « Automatique et télémécanique », t. 25, n° 12, 1964.

7.9. JIVOGLIADOV V.— Sur la commande duale optimale des systèmes à retard pur (en russe). « Automatique et télémécanique », t. 25, n° 1, 1964.

7.10. RESWICK J. B.— Régulateur à retard. Travaux du Ier Congrès international de l'IFAC, t. 1, Editions de l'Académie des Sciences de l'U.R.S.S., 1961. (Cité par Buckley P.S.)

7.11. BUCKLEY P. S.— Régulation automatique des processus à retard pur. Travaux du Ier Congrès international de l'IFAC, t. 1, Editions de l'Académie des Sciences de l'U.R.S.S., 1961.

7.12. JIVOGLIADOV V.— Application de la théorie de la commande duale aux problèmes de commande d'après des paramètres indirects (en russe). Dans le recueil: Théorie et application des systèmes automatiques. Editions Naouka, Moscou, 1964.

7.13. JIVOGLIADOV V.— Théorie des décisions statistiques appliquée à certaines questions d'automatisation des systèmes industriels complexes (en russe). Dans le recueil: Automatisation dans l'industrie des métaux non ferreux. Editions Ilim, Frounzé, 1964.

7.14. JIVOGLIADOV V.— Application de la théorie des décisions statistiques aux problèmes de commande d'après des paramètres indirects (en russe). « Automatique et télémécanique », t, 25, n° 10, 1964.

7.15. SKLANSKY J.— Adaptation and feedback. Joint Automatic Control Conference, New York, Univ., June, 1962.

7.16. GAAZE-RAPOPORT M., FELDBAUM A.— Systèmes auto-adaptatifs (en russe). Encyclopédie d'automatique et d'électronique industrielle, t. III, Editions Sovetskaïa Entsiklopédia, 1964.

7.17. GAMBA A.— Optimum performance of learning machines. Proc. IRE, n° 1961.

7.18. ABRAMSON N., BRAVERMAN D.— Learning to recognize patterns in a random, enviroment. Trans. IRE, PGIT, v. 8, n° 5, 1962.

7.19. BRAVERMAN D.— Learning filters for optimum pattern Recognition. Trans. IRE, PGIT, v. 8, n° 4, July, 1962.

7.20. TSETLINE M.— Sur le comportement des automates finis dans des milieux aléatoires (en russe). « Automatique et télémécanique », t. 22, n° 10, 1961.

7.21. FELDBAUM A.— Sur une classe de systèmes d'auto-apprentissage à commande duale (en russe). « Automatique et télémécanique », t. 25, n° 4, 1964.

INDEX ALPHABÉTIQUE DES MATIÈRES

Accumulation de l'information
— active 28
— indépendante (passive) 28
Algorithme
— de commande 9
d'un filtre 37
— optimal 210
— quasi optimal 421
— d'un récepteur optimal 233
— uniformément optimal 213
Analyse fonctionnelle 108
Automatique 9
Automatisation 424

Bruit
— aléatoire 351
— blanc 56
— de fond 352
But de la commande 20

Calcul variationnel classique 58
Coefficient de corrélation 49
Commande
— admissible 35
— avec facteurs indirects 392
— duale 32
— numérique 17
— par impulsions 17
Conditions de transversalité 69
Consigne 14
Coordonnées suffisantes 260
Covariance 49
Critère
— intégral généralisé 23
— de Kotelnikov 246
— de Neumann-Pearson 247
— d'optimalité 11
— de qualité du processus transitoire 22
Cybernétique
— générale 10
— technique 10
Cycles limites 207

Décision
— de Bayes 247
— du minimax 247
Densité de probabilité
— a posteriori 239
— a priori 249
— bidimensionnelle 51
— conditionnelle 48
— conjointe 51
— inconditionnelle 48
— spectrale 56
— — mutuelle 56
— unidimensionnelle 52
Dérivateur 27
Détecteur synchrone 242

Ecart-type 43
Echantillon d'un processus aléatoire 51
Echantillons 57
Energie
— d'un signal 242
— spécifique 242
Equation
— de Bellman 83
— d'Euler 60
— d'Euler-Lagrange 145
— d'Euler-Poisson 64
— aux variations 94
Equations
— adjointes 89
— — canoniques 91
Erreurs 28
Espace
— du bruit 235
— des décisions 242
— des observations 235
— des paramètres 236
— de phase des erreurs 120
— du risque 247
— du signal 235
Espérance mathématique 25
Estimation suffisante 302
Evénement aléatoire 39

Evénements
— deux à deux indépendants 40
— incompatibles 39
— indépendants 40
Excès de structure 185
Extremum
— absolu 67
— lié 67

Filtre
— linéaire 29
— non linéaire 30
— optimal 29
Fluctuation 126
Fonction
— d'auto-corrélation 54
— de corrélation 38, 54
— — mutuelle 55
— cumulative 42
— de décision 32, 43
— de Dirac 57
— de pénalisation 244
— de pertes 244
— — élémentaire 252
— de pondération 37
— de répartition 42
— scalaire 34
— de vraisemblance 238
Formule de Bayes 41
Fréquence
— d'un événement 40
— — conditionnelle 40

Gain 215
Grandeur
— commandée 13
— de sortie 13
Grossièreté de l'organe de commande 223
Groupe complet 40

Hamiltonien 85
Hypersurface
— de commutation 120
— non stationnaire 120
— quasi stationnaire 120
— stationnaire 121

Identification des images 10
Indétermination 27
Information 26, 424
— complète 28
Information
— courante 29
— incomplète 28, 181
— maximale 181
— redondante 185
— sur le but de la commande 27
— sur la consigne 27
— sur l'opérateur 26
— sur l'état de l'objet 27
— sur la perturbation 26
Intégration cohérente 242
Invariance 225
— à ε près 231
— totale 225
Isosurface 86

Loi
— de commande 31
— de distribution 52
— de Gauss-Laplace 44
— normale 44

Machines de commande 10
Matrice spacio-temporelle 315
Médiane 43
Mesure
— directe 183
— indirecte 184
Méthode
— de compensation 15
— du maximum de vraisemblance 240
— de probabilité a posteriori 239
— de Riesz 69
— variationnelle 58
Minimax 24
Modèle auto-adaptatif optimal 364
Moment
— centré 43
— d'ordre k 43
Moyenne
— d'ensemble 55
— temporelle 55
Multiplicateurs de Lagrange 67

Opérateur de l'objet 16
Optimisateur 10
Organe
— de commande 9
— d'exécution 9
— de mesure 9
— réalisable 223

Pénalité 144
Perturbation 14
— additive 20
— aléatoire 20
— paramétrique 20
Pile chimique 148
Politique 209
Potentiomètre automatique 109
Probabilité 39
— conditionnelle 40
conjointe 51
— d'états 204
— d'un événement composé 52
— inconditionnelle 40
— induite 423
— limite 206
— a posteriori 239
— a priori 238
— de transition 203
— unidimensionnelle 52
Problème
— de Bayes 235
— de Boltz 72
— de commande à extrémité libre de trajectoire 35
— — à point final fixé 36
— dual 225
— de filtration linéaire 37
— de Lagrange 67
— de Mayer 72
— de minimisation du risque 245
— de la recherche de la stratégie optimale 34
— — du processus optimal 34
— de régulation en temps minimal 36
Processus
— aléatoire 51
— — ergodique 55
— — markovien 52
— — non stationnaire 53
— — stationnaire 53
— probabiliste 51
— stochastique 51
Processus aléatoires stationnairement liés 53
Programmation
— dynamique 72
— linéaire 232
Puissance du signal 56

Quantification
— en amplitude 17
— dans le temps 17

Réaction 14
Réalisation d'un processus aléatoire
Récepteur optimal 233
Recherche automatique 10
Rebondance d'information 185
Régulation automatique 10
Risque
— d'action 282
— conditionnel élémentaire 254
— élémentaire 253
— d'étude 282
— moyen 245
— total 245

Schéma équivalent 117
Série de Pika et Silverberg 307
Statistique exhaustive 260
Stratégie 34, 209
— aléatoire 243
— de Bayes 249
déterministe 243
— élémentaire 275
— mixte 249
— optimale 34
— — du minimax 246
— — de la recherche automatique 11
— pure 249
— totale 275
Système
— d'auto-apprentissage 405
— en boucle fermée 14
Système
— en boucle ouverte 14
— continu 17
— continu à états discrets 17
— discret
— — markovien 203
dynamique 223
— idéal 225
— à information maximale sur l'objet 187
— irréductible 297
— neutre 290
— optimal à information complète sur l'objet 28
— optimal à information incomplète sur l'objet 28
— optimal selon le critère du minimax 25
— à paramètres répartis 84
— à programme 286
— à réaction 223
— réductible 270
— à retard pur 383
— sans inertie 184
— sans mémoire 184
— de stabilisation automatique 286

Système
— statistiquement optimal 25
— strictement discret 203
— uniformément optimal 25
Suite
— des commandes 209
— des échantillons 242

Temps de régulation 23
Théorème
— de l'invariance 226
— de Kotelnikov 57
Théorie
— des automates 12
— — finis 12
de compensation 264
— de corrélation 232
— des décisions
— — bialternatives 234
— — multialternatives 234
— — statistiques 232
— d'estimation des paramètres 234
— d'estimation des processus 235
— des fonctions de décision 233
— générale de la commande 10
— générale des signaux 12
— générale des systèmes 12
— de l'information 12 232
Théorie
— des jeux 233
— de stabilité potentielle aux perturbations 233
— des systèmes
— — d'auto-adaptation 10
— — d'auto-apprentissage 10
— — optimaux 12
Trajectoire de phase 36
— — optimale 60, 73

Variable
— aléatoire 39
— — continue 42
Variance 43
Variation 59, 94
— petite 92
— première 59
— seconde 59
Vecteur
— aléatoire
— du bruit 235
— des conditions initiales 25
— du coût 211
— des paramètres 25, 236
— de la perturbation 26
— de sortie 13
— temporel 253

INDEX DES NOMS

Bayes 41, 235
Bellman R. 58, 233
Bernoulli J. 58, 423
Boltianski V. 84
Boltz 72
Borel E. 232
Boutkovski A. 84
Bushaw D. W. 110

Carnap R. 423
Carson 138
Cauchy 159
Cramer H. 232

Desoer 70
Dirac 57, 258
Duhamel 154

Euler 58

Fischer R. 232
Fourier 56

Gamkrélidzé R. 84
Gauss 44

Kantorovitch L. 232
Khintchine A. 38
Kirillova F. 108
Kolmogorov A. 30, 232
Kotelnikov V. 57, 233
Krassovski N. 108
Kreiné M. 108
Kronecker 133
Krotov V. 108
Kulikowski R. 108

Lagrange 58
Laplace 44, 423
Linnik Y. 232

Marianovski D. 109
Markov A. 52
Mayer 72
Middleton D. 233
Miele 70

Neumann J. 232

Pearson E. 232
Pike 307
Poisson 64
Poncelet 226
Pontriaguine L. 58, 84

Reswick 385
Riesz 385, 69
Rozonoère L. 84

Shannon C. 424
Silverberg 307
Smith 385
Svetcharnik D. 109

Taylor 82
Tsypkine I. 7

Viète 66
Voronov 22

Wiener N. 30

À PARAÎTRE

DANS LA MÊME COLLECTION

THÉORIE MATHÉMATIQUE DES PROCESSUS OPTIMAUX

par L. Pontriaguine et coll.

384 pages 14,5×22 cm. Relié toile sous jaquette.

Cet ouvrage de l'académicien L. Pontriaguine contient un exposé détaillé de la théorie des processus optimaux axée sur le principe du maximum. Ce principe fournit des solutions à une série de problèmes mathématiques et pratiques qui, tout en constituant des problèmes de variation, ne peuvent être résolus suivant le schéma classique du calcul des variations.

Le livre s'adresse à des mathématiciens et ingénieurs ayant affaire à des équations différentielles ordinaires, au calcul des variations et à leurs applications. Il sera également fort utile aux étudiants des Universités scientifiques et techniques.

La première édition de cet ouvrage a été couronnée du Prix Lénine.

À PARAÎTRE

DANS LA MÊME COLLECTION

COMMANDE AUTOMATIQUE

par J. Roïtenberg

395 pages 14,5×22 cm. Relié toile sous jaquette.

Ce livre de J. Roïtenberg, professeur à l'Université de Moscou, docteur ès sciences physiques et mathématiques, est consacré à la théorie de la commande automatique dont l'importance de principe s'allie à de nombreuses applications.

Basé sur le cours professé par l'auteur aux étudiants de mécanique de l'Université de Moscou, l'ouvrage expose la théorie des systèmes commandés linéaires unidimensionnels et multidimensionnels, les méthodes de recherches concernant la stabilité et les processus transitoires dans les systèmes linéaires stationnaires, la stabilité absolue des systèmes commandés non linéaires, les systèmes à temps de commande fini, la théorie des fonctions matricielles, les questions de commandabilité et et d'observabilité des systèmes stationnaires et non stationnaires.

Destiné aux étudiants de mécanique, de mathématiques et de physique des établissements d'enseignement supérieur, cet ouvrage rendra également service aux spécialistes des questions de commande.